MATH495 Stochastic Process Note

1 Background

During my exchange in WashU, I've typed some of my notes from lectures and collaborated some with online resources and my understanding. This is mainly for the course - Stochastic Process, as the professor doesn't have a typed note. Since I typed the notes quite rush, there maybe some mistakes or typos. I've typed the note mainly for my revision and probably future review. But I am also glad if this note can help anyone need it.

2 CTMC Basic Properties

Definition 1. *A continuous time process* $\{X_t\}_{t\geq 0}$ *taking values in a countable set* S *is said to be a (continuous-time) markov chain if, for any* $s_0 < s_1 < \cdots < s_n < s < t$ *and* $i, j, i_0, \cdots, i_n$ *we have* $P[X_{t+s} = j | X_s = i, X_{s_n} = i_n, \cdots, X_{s_0} = i_0] = P[X_t = j | X_0 = i]$

Terminology 1. $p_{ij}(t) = P[X_t = j | X_0 = i]$ *are called* ***transition probabilities****. This means the starting state is i when time = 0 and going to j when time = t.*

The matrix $P(t) = \begin{bmatrix} p_{11}(t) & \cdots & p_{1N}(t) \\ \vdots & & \vdots \\ p_{N1}(t) & \cdots & p_{NN}(t) \end{bmatrix}$ *is called the transition probability matrix of the chain.*

Example 1. *A homogeneous Poisson process* $\{N_t\}_{t\geq 0}$ *is a markov process with state space* $S = \{0, 1, 2, \cdots\}$.

Theorem 1. *Suppose* $\{Y_n\}_{n\geq 0}$ *is a discrete-time markov chain and* $\{N_t\}_{t\geq 0}$ *is an independent homogeneous Poisson process. Then,* $X_t = Y_{N_t}$ *is a markov process.*

2.1 Matrix exponential

Lemma 1. $e^A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k$ *where* $A^0 = I$

1. $e^0 = I$ *where 0 is the zero matrix*
2. $e^{cI} = e^c I$ *for some scalar c*
3. $e^{UAU^{-1}} = Ue^A U^{-1}$
4. $AB = BA \Rightarrow e^A e^B = e^{A+B}$

2.2 Fundamental Construction of CTMC

One very large class of CTMC (that cover almost all cases) is based on independent exponential times. The parameters or inputs of the process are some nonnegative constants $q(i,j), i,j \in S$ with $i \neq j$. which are called **jump rates**. (as we shall see $q(i,j)$ is the rate at which the chain jump from i to j).

Next, set

$$\lambda_i = \sum_{j \neq i} q(i,j)$$
$$r_{i,j} = \frac{q(i,j)}{\lambda_i}$$

where λ_i is the rate at which the chain leaves i and $r(i,j)$ is the probability of going to j when leaving i.

2.3 Idea of the Construction Procedure

The previous parameters determine the dynamics of the chain in a simple way. If the chain X arrives to a state i such that $\lambda_i = 0$, then it would stay there forever. But if $\lambda_i > 0$ then the chain will stay there an exponential time d with rate λ. The chain with then decide to jump to another state $j \neq i$ with probability $r(i,j)$.

Equivalently, we can explicitly write $\{X_t\}$ in terms of a discrete time markov chain $\{Y_n\}_{n \geq 0}$ with transition probabilities $r(i,j)$ and a sequence of i.i.d exponential $(\lambda = 1)$ time $\tau_0, \tau_1, \cdots$ as follows:

$$X(t) = Y_n \text{ if } T_n \leq t < T_{n+1}$$

$$\text{where } T_n = t_1 + \cdots + t_n \text{ with } t_i = \frac{\tau_{i-1}}{\lambda(Y_{i-1})} \sim exp(\lambda(Y_{i-1}))$$

Terminology 2. *The discrete-time process $\{Y_n\}_{n \geq 0}$ that records the consecutive states of the chain is called the* ***embedded markov chain*** *and it can be proved that it is indeed a markov chain.*

The rates $q(i,j)$ are typically arrange in the following matrix form

$$Q = \begin{bmatrix} -\lambda_1 & q(1,2) & q(1,3) & \cdots & q(1,N) \\ q(2,1) & -\lambda_2 & q(2,3) & \cdots & q(2,N) \\ \vdots & & & & \vdots \\ q(N,1) & q(N,2) & q(N,3) & \cdots & -\lambda_N \end{bmatrix}$$

This matrix is called ***the infinitesimal generator*** *of the process.*

It is customary to draw the chain as a connected graph with indices representing states and arrows connecting states $i \neq j$ such that $q(i,j) > 0$. Specifically, an arrow goes from i to j if $q(i,j) > 0$. The arrow is labeled by $q(i.j)$. This graph is sometimes called ***transition rate graph***.

2.4 Interpretation of $q(i, j)$

As it turns out, $\lim_{h\to 0} \frac{P[X_{t+h}|X_t=i]}{h} = q(i,j), i \neq j$ or equivalently, $P[X_{t+h} = j|X_t = i] = q(i,j)h + \theta(h)$. The second definition is the probability that a chain in state i jumps to another state j in a small time interval h.

We can see from the definition $q(i, j)$ means the probability goes from state i to state j in a very short time. Therefore, we say this probability as jump rate.

2.5 Limiting distribution

The following theorem formalizes the face that the chain describes in the construction procedure is indeed a CTMC:

Theorem 2. *Let $\{X_t\}t > 0$ be the process. Then $\{X_t\}$ is CTMC with transition probability*

$$P(t) = \begin{bmatrix} P_{11}(t) & \cdots & P_{1N}(t) \\ \vdots & & \vdots \\ P_{N1}(t) & \cdots & P_{NN}(t) \end{bmatrix} = e^{tQ}$$

Lemma 2. *In particular, if Q has the eigendecomposition: $Q = U\Lambda U^{-1}$ where $\Lambda = diag(\alpha_1, \cdots, \alpha_N)$ is the diagonal matrix of eigenvalues, then $P(t) = e^{tQ} = Ue^{t\Lambda}U^{-1}$, where $e^{t\Lambda} = diag(e^{t\alpha_1}, \cdots, e^{t\alpha_N})$.*

Lemma 3. *Note that $\alpha = 0$ is always an eigenvalue for Q corresponding to the eigenvector $[1, 1, \cdots, 1]$. Moreover, note that $\lim_{t\to\infty} P(t)$ exists iff all $\alpha_i \leq 0$.*

Example 2. $Q = \begin{bmatrix} -1 & 1 \\ 2 & -2 \end{bmatrix}$ *Find its transition probabilities and determine its long term limiting distribution, if it exists, $\lim_{t\to\infty} P[X_t = j|X_0 = i] = \pi_j$.*

1. *We first diagonalize Q. Note that we have $\lambda = 0$ or $\lambda = -3$.*
2. *For $\lambda = 0$, the eigenvector $= [1, 1]^T$. For $\lambda = -3$, the eigenvector $= [1, -2]^T$.*
3. *Compute $P_t = Ue^{tQ}U^{-1} = \begin{bmatrix} \frac{2}{3} + \frac{1}{3}e^{-3t} & \frac{1}{3} - \frac{1}{3}e^{-3t} \\ \frac{2}{3} - \frac{2}{3}e^{-3t} & \frac{1}{3} + \frac{2}{3}e^{-3t} \end{bmatrix}$*
4. $\lim_{t\to\infty} P(t) = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}$ *so $\pi = [\frac{2}{3}, \frac{1}{3}]$*

Remark. *In the example above, $[\frac{2}{3}, \frac{1}{3}]$ is the limiting distribution. This coincides with the row of U^{-1} corresponding to the eigenvalue $\alpha = 0$. In that example, the top row of U^{-1}. The limit $\lim_{t\to\infty} P[X_t = i]$ exists whenever $\alpha_i \leq 0, \forall i$.*

But, in order for this not to depend on the initial distribution, we must have that 0 is a simple eigenvalue and all other α_i must be negative.

3 CTMC Absorption times Probability

For a discrete-time MC¡ we consider the following problems:

1. **Absorption Times Problems**

   The average time it takes for the chain to be absorbed by one of the recurrent classes when starting at a transient state.

2. **Absorption Probability Problems**

   What is the probability that when starting at some state i, the chain will be absorbed by a specific recurrent class instead of others?

In this section, we will show how to answer this kind of questions for a CTMC.

3.1 Solution procedure for problem 1

For the absorption times problem, suppose A consists of all recurrent states and $T_r = S \setminus A$ are the transient states. We want to compute:

$$g(i) = E_i[T_A] \text{ where } T_A = inf\{t \geq 0 : X_t \in A\}$$

In this case, we have $g = \begin{bmatrix} g(i_1) \\ \vdots \\ g(i_m) \end{bmatrix} = (-Q^{sub}) \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$ where Q^{sub} is the submatrix of Q corresponding to the transient states.

Example 3. *A shop has two barbers that can cut hair at rate 3 people per hour. Customers arrive at times of a rate 2 Poisson process, but will leave if there are two people getting their haircut and two waiting. The state of the system to be the number of people in the shop. Find $P_i[V_0 < V_4]$ for $i = 1, 2, 3$.*

(a) Derive the infinitesimal generator Q.
(b) Compute $E[T_0|X_0 = x]$, $x = 1, 2, 3, 4$, where $T_0 = inf\{t > 0 : X_t = 0\}$.

We can find the transition rate graph first with $q(i, i+1) = 2$ for $i = 0, 1, 2, 3$, $q(i, i-1) = 6$ for $i = 2, 3, 4$ and $q(1, 0) = 3$. $q(i, i+1) = 2$ because customers arrive at times of a rate 2. We have two barbers so the shop can have 6 people hair cut per hour. Therefore, when we have more than 1 customer, $q(i, i-1) = 6$. While if we only have 1 customer, then it is not related to the number of the barbers so $q(1, 0) = 3$.

Hence, we can write $Q = \begin{bmatrix} -2 & 2 & 0 & 0 & 0 \\ 3 & -5 & 2 & 0 & 0 \\ 0 & 6 & -8 & 2 & 0 \\ 0 & 0 & 6 & -8 & 2 \\ 0 & 0 & 0 & 6 & -6 \end{bmatrix}$.

Note that $q(0,1) = 2$ means the rate of arrival of new customers; $q(1,0) = 3$ means the rate at which 1 customer is served by the barber; $q(2,1) = 6$ means the rate at which the two customers will be served.

Suppose $X_0 = 1$ (only one customer in the shop). Let D_1 be the time to serve the customer $\sim exp(\lambda = 3)$ and A_1 be the arrival time of a new customer $\sim exp(\lambda = 2)$. Then, the time the chain will spend in time 0 is $T_0 = min\{D_1, A_1\} \sim (3+2) \Rightarrow \lambda_0 =$ holding rate at time 0 $= 5$. After T_0, the chain will jump to either state 0 or state 2 with probabilities $r(1,0) = P[D_1 < A_1] = \frac{3}{5}$ and $r(1,0) = P[D_1 > A_1] = \frac{2}{5}$. Recalling that $r(i,j) = \frac{q(i,j)}{\lambda_i}$, we get the rates $q(1,0) = \frac{3}{5} \times 5 = 3$ and $q(1,2) = \frac{2}{5} \times 5 = 2$

Similarly, if $X_0 = 2$, the time X spends in time 2 is $T_2 = min\{D_1, D_2, A_1\} \sim exp(3 + 3 + 2)$. Then $\lambda_2 = 8$. After T_2, the chain jumps to state 1 or 3 with probs: $r(2,1) = P(min\{D_1, D_2\} < A_1) = \frac{6}{8}$ and $r(2,3) = P(min\{D_1, D_2\} > A_1) = \frac{2}{8}$. Then, we get the rates $q(2,1) = \frac{6}{8} \times 8 = 6$ and $q(2,3) = \frac{2}{8} \times 8 = 2$.

Now, we discuss part b. The idea is to think of state 0 as an absorbing state and then follow the procedure.

1. *Extract the matrix Q^{sub}.*

$$Q^{sub} = \begin{bmatrix} -5 & 2 & 0 & 0 \\ 6 & -8 & 2 & 0 \\ 0 & 6 & -8 & 2 \\ 0 & 0 & 6 & -6 \end{bmatrix}.$$

2. *Compute $(-Q^{sub})^{-1}$*

3. *Multiply by $(1,1,1,1)^T$*

Hence, we get $g(i) = (\frac{40}{81}, \frac{119}{162}, \frac{155}{162}, \frac{91}{81})$ where $E[T_0|X_0 = 0] = \frac{40}{81}$.

3.2 Solution procedure for problem 2

For absorption probability, it should be clear that we can consider the analogous question for the embedded discrete-time chain $\{Y_n\}_{n \geq 0}$ with transition probabilities $r(i,j) = \frac{q(i,j)}{\lambda_i}$.

But we can also use directly Q. Concretely, suppose we have 2 closed recurrent classes R_1 and R_2 and the rest are transient $T_r = S \setminus (R_1 \cup R_2) = i_1, \cdots, i_m$.

We want to compute $l(i) = P_i[T_{R_1} < T_{R_2}]$, $i \in T_r$. Then $\begin{bmatrix} l(i_1) \\ \vdots \\ l(i_m) \end{bmatrix} = (-Q^{sup})^{-1} \begin{bmatrix} V_{R_1}(i_1) \\ \vdots \\ V_{R_1}(i_m) \end{bmatrix}$

where Q^{sup} is the submatrix of Q corresponding to the transient states and

$$V_{R_1} = \begin{bmatrix} V_{R_1}(i_1) \\ \vdots \\ V_{R_1}(i_m) \end{bmatrix} = \begin{bmatrix} \sum_{j \in R_1} q(i_1, j) \\ \vdots \\ \sum_{j \in R_1} q(i_m, j) \end{bmatrix}.$$

Example 4. *A shop has two barbers that can cut hair at rate 3 people per hour. Customers arrive at times of a rate 2 Poisson process, but will leave if there are two people getting their haircut and two waiting. The state of the system to be the number of people in the shop. Find $P_i[T_0 < T_4]$ for $i = 1, 2, 3$.*

The idea is to treat 0 and 4 as absorbing states and following the below procedure:

1. *Extract the matrix Q^{sub}*

$$Q^{sub} = \begin{bmatrix} -5 & 2 & 0 \\ 6 & -8 & 2 \\ 0 & 6 & -8 \end{bmatrix}.$$

2. *Compute $(-Q^{sup})^{-1}$*

3. *Compute* $V_{R_1} = \begin{bmatrix} q(1,0) \\ q(2,0) \\ q(3,0) \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}.$

4. *Compute* $h = (-Q^{sup})^{-1} V_{R_1} = \begin{bmatrix} \frac{39}{41} \\ \frac{36}{41} \\ \frac{27}{41} \end{bmatrix}.$

So, for instance, when having 1 initial customer, about 95% of the time the shop will eventually be empty before being full.

4 Final Remarks

4.1 Limiting Behaviour

As in the DTMC, we define the limiting distribution of a CTMC as a vector $\tilde{\pi}$ such that

$$\lim_{t \to \infty} P_{i,j}(t) = \lim_{t \to \infty} P(X_t = j | X_0 = i) = \tilde{\pi}_j$$

for all j, regardless of i.

This is equivalent to either of the following

1. $\lim_{t \to \infty} \alpha P(t) = \tilde{\pi}$, for any initial distribution α

2. $\lim_{t\to\infty} P(t) = \begin{bmatrix} \tilde{\pi} \\ \vdots \\ \tilde{\pi} \end{bmatrix}$

4.2 Existence of the long-term limiting distributions

We have already seen one criterion for the existence of the long-term limiting distributions. Namely, if all conditions below hold true:

1. Q is diagonalizable.
2. The eigenvalue $\lambda = 0$ is simple.
3. All nonzero eigenvalues are negative.

In fact, it can be shown that π is the row corresponding to $\lambda = 1$. For example, if the diagonal 1 is in the last row, then π is the last row of U^{-1}.

4.3 Ergodic Theorem

Is there another criterion in terms of dynamical concepts such as irreducibility? Yes. The main theorem is the following:

Theorem 3. *Suppose that a continuous Markov chain is irreducible (i.e., its embedded Markov chain is irreducible) and has a stationary distribution π, then*

$$\lim_{t\to\infty} P(X_t = j | X_0 = i) = \pi_j.$$

Remark. *The embedded Markov chain Y is irreducible if $i \mapsto j$, for any $i \neq j$. Recall that $r(i,j) = \frac{q(i,j)}{\lambda_i}$ is the transition probability matrix of Y. Thus, the chain is irreducible if, for any $i \neq j$. There is a path that goes from i to j in the transition rates graph of the chain.*

5 Brownian Motion

Brownian motion is a stochastic process, which is rooted in a physical phenomenon discovered almost 200 years ago. In 1827, the botanist Robert Brown, observing pollen grains suspended in water, noted the erratic and continuous movement of tiny particles ejected from the grains. He studied the phenomenon for many years, ruled out the belief that it emanated from some "life force" within the pollen, but could not explain the motion. Neither could any other scientist of the 19th century.

In 1905, Albert Einstein solved the riddle in his paper, On the movement of small particles suspended in a stationary liquid demanded by the molecular-kinetic theory of heat. Einstein explained the movement by the continual bombardment of the immersed particles by the molecules in the liquid, resulting in "motions of such magnitude that these motions can easily be detected by a microscope". Einstein's theoretical explanation was confirmed 3 years later by empirical experiment, which led to the acceptance of the atomice nature of matter.

5.1 Standard Brownian motion or Wiener Process

Definition 2. *A continous-time process* $\{B_t\}t \geq 0$ *is a standard Brownian motion (BM) if it satisfies the following properties:*

1. $B_0 = 0$ *(starts at time = 0)*
2. $B_{t+s} - Bs \sim N(0, t)$, *for any* $t > 0, s \geq 0$ ***(Stationary increments)***
3. $B_{t_1} - B_{t_0}, B_{t_2} - B_{t_1}, \cdots, B_{t_n} - B_{t_{n-1}}$ *are independent for any* $0 \leq t_0 \leq t_1 \leq \cdots \leq t_n$ ***(Independent increments)***
4. $t \mapsto B_t$ *is a continuous finction of* t ***(Continuous paths)***

5.1.1 Other closely related processes

A process $\{X_t\}_t \geq 0$ is a **BM with variance** σ^2 if it satisfies (1), (3) and (4) but (2) is replaced by $B_{t+s} - Bs \sim N(0, \sigma^2 t)$, for any $t > 0, s \geq 0$.

In addition, if (1) is replaced by $X_0 = x$, then we say that X is a **BM started at x with variance** σ^2. Note that $X_t = x + \sigma B_t$, where B is a standard BM.

Unless otherwise stated, a BM refers to a standard Brownian motion.

Remark.

1. $B_t \sim N(0, t) \Rightarrow E[B_t] = 0$ *and* $Var(B_t) = E[B_t^2] = t$
2. B_s *and* $B_{s+t} - B_s$ *are independent for any* $t, s \geq 0$.
3. *This definition is very similar to that of Poisson process but instead of a Poisson distribution, we have a Gaussian distribution.*
4. $B_{s+t} - B_s \doteq B_t \doteq \sqrt{t}B_1 \doteq \sqrt{t}Z$, *where* $Z \sim N(0, 1)$ ***Again, they have the same distribution but they are not equal.***

Example 5. *A particle's position is modeled with a standard BM. If the particles is at position 1 at time* $t = 2$*, find the probability that its position is at most 3 at time* $t = 5$*.*

$$
\begin{aligned}
P(B_5 \leq 3|B_2 = 1) &= P(B_5 - B_2 \leq 3 - B_2|B_2 = 1) \\
&= P(B_5 - B_2 \leq 2|B_2 = 1) \\
&= P(B_5 - B_2 \leq 2) \\
&= P(Z \leq \frac{2}{\sqrt{3}}) \\
&= 0.876
\end{aligned}
$$

Note the third equality follows by the independence of $B_5 - B_2$ and B_2. The second last equality follows because $B_5 - B_2 \sim N(0, 5 - 2)$.

Example 6. *Find the covariance $Cov(B_s, B_t)$, $E[B_sB_t]$ and the correlation $Corr(B_s, B_t)$.*

$$
\begin{aligned}
E[B_sB_t] &= E[B_s(B_t - B_s + B_s)] \\
&= E[B_s(B_t - B_s)] + E[B_s^2] \\
&= E[B_s]E[B_t - B_s] + Var(B_s) \\
&= s
\end{aligned}
$$

The third equality follows by the independent increment and the last equality is because $E[B_s] = 0$ and $Var(B_s) = s$.

$$
\begin{aligned}
Corr(B_s, B_t) &= \frac{Cor(B_s, B_t)}{\sqrt{Var(B_s Var(B_t)}} \\
&= \frac{min\{s, t\}}{\sqrt{st}}
\end{aligned}
$$

Example 7. *Find the distribution of $B_s + B_t$.*

Assume $s \leq t$. We have $B_s + B_t = 2B_s + (B_t - B_s) \doteq N(0, 3s + t)$.

The equality follows because $2B_s \sim N(0, 4s)$ and $B_t - B_s \sim N(0, t - s)$.

5.1.2 Gaussian Process

Theorem 4. *The standard BM $\{B_t\}_{t \geq 0}$ is such that,* ***for any sequence of constant times*** $0 < t_1 < \cdots < t_n$***, and any constant*** $a_1, \ldots, a_n \in \mathbb{R}$***,*** $a_1B_{t_1} + \cdots + a_nB_{t_n}$ ***is normally distributed.*** *Equivalently, for any $t_1 < t_2 < \cdots < t_n$, $(B_{t_1}, \ldots, B_{t_n})$ are jointly normally distributed (or has a multivariate normal distribution).*

Definition 3. *A process satisfying the above theorem is called a* ***Gaussian Process****.*

Theorem 5. *A process $\{B_t\}_{t\geq 0}$ is a standard BM if and only if it is a Gaussian process with the following properties:*

1. $B_0 = 0$
2. $E[B_t] = 0$, $\forall t$ ***(mean function)***
3. $Cov(B_s, B_t) = min\{s, t\}$ ***(covariance function)***
4. $t \mapsto B_t$ *is continuous*

Example 8. *Compute $P(B_{1.5} - 2B_2 + B_3 > 2)$.*

Since we already know that $B_{1.5} - 2B_2 + B_3$ is normally distributed, we only need to find its mean and variance. Denote $Y := B_{1.5} - 2B_2 + B_3$. Note that we have $E[Y] = 0$ and $Var(Y) = Var(B_{1.5}) + 4Var(B_2) + Var(B_3) - 4Cov(B_{1.5}, B_2) + 2Cov(B_{1.5}, 3) - 4Cov(B_2, B_3) = 1.5$ Then, $P(Y) = 1 - \Phi(\frac{2}{\sqrt{1.5}})$.

5.1.3 Conditional Brownian Motion

Theorem 6. *For any $0 < t < u$,*

1. $B_u | B_t \sim N(B_t, (u - t))$
2. $B_t | B_u \sim N(\frac{t}{u} B_u, \frac{t(u-t)}{u})$

Remark. *Both conditional properties in the previous theorem have stronger versions. Broadly, for any $0 \leq t < u$,*

1. $B_u | \{B_s\}_{s \leq t} \sim N(B_t, u - t)$
2. $B_t | \{B_s\}_{s \leq t} \sim N(\frac{t}{u} B_u, \frac{t(u-t)}{u})$

Some ideas are as follows:

1. $E[B_t|B_s] = E[B_t - B_s + B_s|B_s] = E[B_t - B_s|B_s] + E[B_s|B_s] = E[B_t - B_s] + B_s = 0 + B_s = B_s$
2. We want $E[B_t|B_s]$ and $Var(B_t|B_s)$. Note that B_t and $B_s - \frac{s}{t} B_t$ are independent. Since $Cov(B_t, B_s - \frac{s}{t} B_t) = Cov(B_t, B_s) - \frac{s}{t} Cov(B_s, B_t) = s - \frac{s}{t} \times t = 0$

Theorem 7. *BM is a markov process*

$$P(B_{s+t} \leq y | B_s = x) = P(B_{s+t} - B_s \leq y - x | B_s = x) = P(B_{s+t} - B_s \leq y - x) = \Phi(\frac{y-x}{\sqrt{t}})$$

It follows that $B_{s+t}|B_s \sim N(B_s, t)$

Theorem 8. *For any fixed time $s \in (0, \infty)$, TFAE*

1. $B_{t+s}|B_s \sim N(B_s, t)$

2. $B_{t+s} - B_s|B_s \sim N(0, t)$

3. $\{B_{t+s}\}_{t\geq 0}$ *is a BM independent of* $\{B_u\}_{u\leq s}$

5.2 Markov property of Brownian Motion

We already saw a version of the markov property for a B.M. $\{B_t\}_{t\geq 0}$: For any $s, t \geq 0$, $B_{s+t}|B_s \sim N(B_s, t)$.

We can write this in the following way: $P(B_{s+t} \leq y|B_v, v \leq s) = P(B_{s+t} \leq y|B_s) = \Phi(\frac{y-B_s}{\sqrt{t}})$ This equality means that, for any $0 \leq t_1 < \cdots < t_n = s$, $P(B_{s+t} \leq y|B_s, B_{t_{n-1}} \cdots B_{t_1}) = P(B_{s+t} \leq y|B_s)$ which now looks almost the same as the Markov Property for CTMC.

5.2.1 Simple Markov Property

Theorem 9. ***(Simple Markov Property)*** *For any fixed time* $s > 0$*, the process* $\{B_{s+t} - B_s\}_{t\geq 0}$ *is a standard BM and this is also independent of the path or history of BM up to time s,* $\{B_u\}_{0\leq u\leq s}$*.*

Therefore, not only $B_{t+s} - B_s \sim N(0, t)$, independent of $\{B_u\}_{u\leq s}$, but furthermore, the whole path $\{B_{t+s} - B_s\}_{t\geq 0}$ is a standard B.M., independent of $\{B_u\}_{u\leq s}$.

Thus, e.g., for any $t_1 < t_2 < \cdots < t_n$ and $s_1 < s_2 < \cdots < s_m < s$, we have

$$\begin{aligned} &P(B_{t_1} - B_s \leq x_1, \cdots, B_{t_n} - B_s \leq x_n|B_s, B_{s_1}, \cdots, B_{s_m}) \\ =&P(B_{t_1} \leq x_1, \cdots, B_{t_n} \leq x_n) \end{aligned}$$

We can further show that s can be replace with a random time satisfying certain conditions. The most important case is when s is replace with a "hitting time"

Definition 4. $T_a = min\{t > 0 : B_t = a\}$ *for* $a \in \mathbb{R}$*.*

5.2.2 Strong Markov Property for Hitting Time

Let $a \in \mathbb{R}$ and $T_a = min\{t > 0|B_t = a\}$. Then the translated B.M. $\{B_{t+T_a} - B_{T_a}\}_{t\geq 0}$ is a BM independent of $\{B_u\}_{u\leq T_a}$. It means that the translated B.M. starts from (T_a, a), i.e. we move the origin to (T_a, a) and restarts everything, while the past history can't affect the future motion.

Theorem 10. *Note that we have* $P(T_a < t) = 2P(B_t > a) = 2(1 - \Phi(\frac{a}{\sqrt{t}}))$*.*
In particular, the pdf of T_a *is given by* $f_{T_a}(t) = \frac{a}{\sqrt{2\pi t^3}} e^{-\frac{a^2}{2t}}$*, for* $t > 0$*.*

Remark. *The strong Markov property is used to find the distribution of T_a. Consider standard Brownian motion. At any time t, B_t is equally likely to be above or below the line $y = 0$. Assume that $a > 0$. For Brownian motion started at a, the process is equally likely to be above or below the line $y = a$. This gives, $P(B_t > a|T_a < t) = \frac{1}{2}$.*

Example 9. *A particle moves according to Brownian motion started at $x = 1$. After $t = 3$ hours, the particle is at level 1.5. Find the probability that the particle reaches level 2 sometimes in the next hour.*

For $t \geq 3$, the translated process is a Brownian motion started at $x = 1.5$. The event that the translated process reaches level 2 in the next hour, is equal to the event that a standard Brownian motion first hits level $a = 2 - 1.5 = 0.5$ in $[0, 1]$. The desired probability is:

$$\begin{aligned} P(max\{1 + B_u\} \geq 2, 3 \leq u \leq 4|1 + B_3 = 1.5) &= P(max\{B_u - B_3\} \geq 0.5|B_3 = 0.5) \\ &= P(B_{3+t} - B_3 \geq 0.5, t \leq 1) \\ &= P(B_t \geq 0.5, t \leq 1) \\ &= P(T_{0.5} < 1) \end{aligned}$$

Note that $1+B_u$ means the B.M. started at 1. We subtract $1+B_3 = 1.5$ from $max\{1+B_u\} \geq 2, 3 \leq u \leq 4$ in first equality. Then we use simple markov property in the second and third equality. By the above theorem, $P(T_{0.5} < 1) = 2(1 - \Phi(\frac{1}{2})) = 0.617$.

Example 10. *A laboratory instrument takes annual temperature measurements. Measurement errors are assumed to be independent and normally distributed. As precision decreases over time, errors are modeled as standard Brownian motion. For how many years can the lab be guaranteed that there is at least 90% probability that all errors are less than 4 degrees?*

The problem asks for the largest t such that $P(M_t \leq 4) \geq 0.90$. We have

$$0.90 \leq P(M_t \leq 4) = 1 - P(M_t > 4) = 1 - 2P(B_t > 4) = 2P(B_t \leq 4) - 1.$$

This gives $0.95 \leq P(B_t \leq 4) = P(Z \leq \frac{4}{\sqrt{t}})$, where Z is a standard normal random variable. The 95th percentile of the standard normal distribution is 1.645. Solving $4/\sqrt{t} = 1.645$ gives

$$t = (\tfrac{4}{1.645})^2 = 5.91 \text{ years.}$$

5.2.3 Zeros of Brownian Motion

Theorem 11. *For $0 \leq r < t$, let $z_{r,t}$ be the probability that standard Brownian motion has at least one zero in (r, t). Then,*

$$z_{r,t} = \tfrac{2}{\pi} arccos(\sqrt{\tfrac{r}{t}})$$

This is probably be tested in the exam paper so I also include the proof here.
Assume that $B_r = x < 0$. The probability that $B_s = 0$ for some $s \in (r,t)$ is equal to the probability that for the process started in x, the maximum on $(0, t-r)$ is greater than 0. By translation, the latter is equal to the probability that for the process started in 0, the maximum on $(0, t-r)$ is greater than $|x|$. That is,

$$P(B_s = 0 \text{ for some s} \in (r,t)|B_r = x) = P(M_{t-r} > |x|)$$

For $x > 0$, consider the reflected process $-B_s$ started in $-x$. In either case,

$$\begin{aligned}
z_{r,t} &= \int_{-\infty}^{\infty} P(M_{t-r} > |x|) \frac{1}{\sqrt{2\pi r}} e^{-x^2/2r} dx \\
&= \int_{-\infty}^{\infty} \int_0^{t-r} \frac{|x| e^{-x^2/2s}}{\sqrt{2\pi s^3}} ds \frac{1}{\sqrt{2\pi r}} e^{-x^2/2r} dx \\
&= \frac{1}{\pi} \int_0^{t-r} \frac{1}{\sqrt{rs^3}} \int_0^{\infty} x e^{-x^2(r+s)/2rs} dx ds \\
&= \frac{1}{\pi} \int_0^{t-r} \frac{1}{\sqrt{rs^3}} \int_0^{\infty} e^{-z(r+s)/rs} dz ds \\
&= \frac{1}{\pi} \int_0^{t-r} \frac{1}{\sqrt{rs^3}} \frac{rs}{r+s} ds \\
&= \frac{1}{\pi} \int_{r/t}^{1} \frac{1}{\sqrt{x(1-x)}} dx \\
&= \frac{2}{\pi} (arcsin(\sqrt{1}) - arcsin(\sqrt{r/t})) \\
&= \frac{2}{\pi} arccos(\sqrt{\frac{r}{t}})
\end{aligned}$$

5.2.4 Last Zero Standing

Theorem 12. *Let L_t be the last zero in $(0,t)$. Then,*

$$P(L_t \leq x) = \tfrac{2}{\pi} arcsin(\sqrt{\tfrac{x}{t}}), \textit{ for } 0 < x < t.$$

Remark. *The density of L_t is called arc-sine density and is given by:*

$$f_{L_t}(x) = \frac{\mathrm{d}P(L_t \leq x)}{\mathrm{d}x} = \frac{\mathrm{d}}{\mathrm{d}x}(\tfrac{2}{\pi} arcsin(\sqrt{\tfrac{x}{t}})) = \frac{\sqrt{t}}{\pi\sqrt{x(t-x)}},\ x \in (0,t).$$

5.2.5 The Reflection Trick or Principle

This is another way to look into the identity $P(T_a < t) = 2P(B_t > a)$. Consider a path that hits level a before time t. Reflect the path after the hitting time T_a and consider the resulting process $\tilde{B}$:

$$\tilde{B} = \begin{cases} B_t & \text{if } 0 \le t \le T_a \\ 2a - B_t & \text{if } t \ge T_a \end{cases}$$

Intuitively, for every path of B that hits a before t and ends up above a at time t, there is one path that hits a before t and ends up below a at time t, and vice versa. So, it makes sense that

$$P(T_a < t, B_t > a) = P(T_a < t, B_t < a)$$

Finally, we have the following:

$$\begin{aligned} P(T_a < t) &= P(T_a < t, B_t > a) + P(T_a < t, B_t < a) \\ &= 2P(T_a < t, B_t > a) \\ &= 2P(B_t > a) \end{aligned}$$

Note that the second equality follows by $\{B_t > a\} \subseteq \{T_a < t\}$.

5.2.6 Hitting Time applications summary

1. Different versions of the Reflection Principle:
   (a) $a > 0 : P(max_{s \le t}\{B_s \ge a\}) = 2P(B_t \ge a) = P(T_a \le t)$, where $T_a = inf\{t \ge 0 : B_t = a\}$
   (b) $\max_{s \le t} B_s \doteq |B_t|$
   (c) $P(T_a < t, B_t > a) = P(T_a < t, B_t < a)$
2. Zeros of B.M

   For $0 \le r < t$, $P(B_s = 0, s \in (r,t)) = \frac{2}{\pi} arccos(\sqrt{\frac{r}{t}})$
3. Last Zero Standing

   $L_t = max\{0 \le s \le t : B_s = 0\}$

   $P(L_t \le x) = \frac{2}{\pi} arcsin(\sqrt{\frac{x}{t}}), 0 < x < t$

5.2.7 Geometric Brownian Motion

Definition 5. *Let $\{X_t\}_{t \ge 0}$ be a Brownian motion with drift parameter μ and variance parameter σ^2. The process $\{G_t\}_{t \ge 0}$ defined by $G_t = G_0 e^{X_t}$ for $t \ge 0$, where $G_0 > 0$, is called geometric Brownian motion.*

Remark. *1.* $E[G_t] = G_0 e^{t(\mu + \sigma^2/2)}$

2. $Var(G_t) = G_0^2 e^{2t(\mu+\sigma^2/2)}(e^{t\sigma^2} - 1)$

3. $E[ln(G_t)] = ln(G_0) + \mu t$

4. $Var(ln(G_t)) = \sigma^2 t$

5.3 Brownian Martingales

The term martingale comes from the theory of game of chance and means a "fair" game where it is equally likely to win or lose in each bet.

Roughly, a martingale is a process whose future value tend to remain constant even if the part history or path of the process is known.

5.3.1 Introduction to Martingales

Definition 6. *A process $\{X_t\}_{t\in\tau}$ (in discrete or continuous time) with time indices τ is a martingale if*

1. $E[|X_t|] < \infty$*, for all* $t \in \tau$,

2. $E[X_t|X_u, u \leq s] = X_s$*, for any* $s < t(s, t \in \tau)$.

Remark. *Condition 2 is equivalent to the following (which is usually how we check it):*

For any "past" times $u_1 < u_2 < \cdots < u_m < s$,

$$E[X_t|X_s, X_{u_1}, \cdots, X_{u_m}] = X_s.$$

In discrete-time $\tau = 0, 1, 2, \cdots$*, condition 2 is equivalent to*

$$E[X_{n+1}|X_n, X_{n-1}, \cdots, X_0] = X_n, \text{ for any } n = 0, 1, \cdots$$

Example 11. ***(Brownian motion)*** *We saw that* $E[B_t|B_u, u \leq s] = B_s$

Example 12. ***(Symmetric Random Walks)*** $X_t = \sum_{i=1}^{t} Y_i$, $t \in \tau = \mathbb{N} = \{0, 1, \cdots\}$, *where* Y_i *are i.i.d with* $P(Y_i = \pm 1) = \frac{1}{2}$.

Example 13. ***(Compensated Poisson Process)*** $X_t = N_t - \lambda t$*, where* $\{N_t\}$ *is a Poisson process with intensity* λ.

5.3.2 Generalized Martingales

Definition 7. *We say that a process $\{X_t\}_{t\in\tau}$ is a martingale with respect to another process $\{Y_t\}_{t\in\tau}$ if*

1. $E[|X_t|] < \infty$*, for all* $t \in \tau$,

2. $E[X_t|Y_u, u \le s] = X_s$, for any $s < t(s,t \in \tau)$.

To understand martingales, we can imagine it is a line with some variations in different time. But for long enough time, we can see the line as a constant line. For example, if we have 100 dollars before we enter casino, we are expected to have 100 dollars when we leave. Note the value we are talking here is the expectation but not the exact value. To better understand and use of martingales, readers are recommend to take measure theory.

Example 14. *Let $X_t = B_t^2 - t$. Then, $\{X_t\}_{t\ge0}$ is a martingale with respect to $\{B_t\}_{t\ge0}$.*

The following gives a whole class of martingales related to Brownian motion.

Theorem 13. *Heat Equation*
If $u(t,x) : [0,\infty) \times \mathbb{R} \mapsto \mathbb{R}$ is smooth such that

$$\tfrac{\partial u}{\partial t} + \tfrac{1}{2}\tfrac{\partial^2 u}{\partial x^2} = 0$$

then $X_t = u(t, B_t)$ is a martingale w.r.t B.

5.3.3 Key Property of Martingales

The most important property of martingales is that the expectation is constant if no effective information is given. i.e. we can conclude the following result:

$$E[X_t] = E[X_0], \text{ for any } t \in \tau$$

Example 15. *Using the property above with the martingale $X_t = e^{\sigma B_t - \frac{\sigma^2}{2}t}$, compute $E[e^{\sigma B_t}]$.*

By property, we have $E[e^{\sigma B_t - \frac{\sigma^2}{2}t}] = E[e^{\sigma B_0 - \frac{\sigma^2}{2}0}] = 1$. But the LHS is just $E[e^{\sigma B_t}]e^{\frac{-\sigma^2}{2}t}$. Then, $E[e^{\sigma B_t}] = e^{\frac{\sigma^2}{2}t}$

5.3.4 Optional Stopping Theorem

The following is one of the main theorem related to martingales.

Theorem 14. *(Optional Stopping Theorem)*
Let $\{X_t\}_{t\ge0}$ be a martingale with respect to a stochastic process $\{Y_t\}_{t\ge0}$. Assume that T is a stopping time for $\{Y_t\}_{t\ge0}$. Then, $E[X_T] = E[X_0]$ if one of the followings are satisfied.

1. *T is bounded. That is, $T \le c$ for some constant c.*
2. *$P(T < \infty) = 1$ and $E[|X_t|] \le c$, for some constant c, whenever $T > t$.*

Remark. *In discrete-time, we also have $E[X_T] = E[X_0]$ where T takes values on $\tau = \{0, 1, \cdots\}$ and the property of 'stopping time' reduces to asking that, for any n, the event $\{T = n\}$ depends on $Y_0, Y_1, \cdots, Y_n$.*

The most important type of stopping time is a hitting time T_a and its combinations: $min\{T_a, T_b\}$, $max\{T_a, T_b\}$, etc.

Example 16. *Let $a, b > 0$. For a standard Brownian motion, find the probability that the process hits level a before hitting level $-b$.*

Let p be the desired probability. Consider the time T that Brownian motion first hits either a or $-b$. That is, $T = min\{t : B_t = a \text{ or } B_t = -b\}$. Observe that $B_T = a$, with probability p, and $B_t = -b$, with probability $1-p$. By the optional stopping theorem, (applied to the martingale $X_t = B_t$)

$$0 = E[B_0] = E[B_T] = pa + (1-p)(-b).$$

Solving for p gives $p = \frac{b}{a+b}$.

Example 17. ***(Expected hitting time)***
Apply the optional stopping theorem with the same stopping time as in the previous example, but with the quadrtic martingale $B_t^2 - t$. This gives

$$E[B_T^2 - T] = E[B_0^2 - 0] = 0,$$

from which it follows that

$$E[T] = E[B_T^2] = a^2 \tfrac{b}{a+b} + b^2 \tfrac{a}{a+b} = ab$$

Note that

$$B_t = \begin{cases} a & \text{with probability } p = \frac{a}{a+b} \\ -b & \text{with probability } 1-p = \frac{b}{a+b} \end{cases}$$

We have thus discovered that the expected time that standard Brownian motion first hits the boundary of the region defined by the lines $y = a$ and $y = -b$ is ab.

5.4 Revision Examples

Example 18. *A facility has four machines, with two repair workers to maintain them. Individual machines fail on average every 10 hours. It takes an individual repair worker on average 4 hours to fix a machine. Repair and failure times are independent and exponentially distributed. Let $X_t \in \{0, 1, 2, 3, 4\}$ be the number of machines working at time t.*

1. *Describe the duration time the chain X stays on each state and the transition probabilities of jumping from each state to other.*
2. *Find the generator matrix Q of the chain.*
3. *Find the stationary distribution.*
4. *In the long term, how many machines are typically operational?*
5. *If all four machines are initially working, find the probability that only two machines are working after 5 hours.*

Though not explicitly said, as customary, all times involves in this problem are independent and exponentially distributed. Suppose at time 0, $X_0 = 0$ (none of the machines are working). The chain will move from 0 to 1 machine working when one of the two repairmen finish. This time is the minimum of the two repair times, so exponential with rate $\lambda_0 = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$.

Suppose there are only one machine working. The chain will remain there until $T_1 = min\{T_{R_1}, T_{R_2}, T_F\} \sim exp(\lambda_1 = \frac{1}{4} + \frac{1}{4} + \frac{1}{10})$, where T_{R_i} = time it takes repairman i to finish and T_F = time it takes for the working machine to fail. Then, the chain jumps to 2 with probability

$$r(1,2) = P(min\{T_1, T_2\} < T_F) = \frac{2/4}{2/4+1/10} = \frac{5}{6}$$

or jumps to 0 with probability

$$r(1,0) = P(T_F < min\{T_1, T_2\}) = \frac{1/10}{2/4+1/10} = \frac{1}{6}.$$

Suppose there are two machines working. The chain will remain there until $T_2 = min\{T_{R_1}, T_{R_2}, T_{F_1}, T_{F_2}\} \sim expo(\lambda_2 = \frac{2}{4} + \frac{2}{10})$. Then, the chain jumps to 3 with probability

$$r(2,3) = P(min\{T_{R_1}, T_{R_2}\} < min\{T_{F_1}, T_{F_2}\}) = \frac{2/4}{2/4+2/10} = \frac{5}{7}$$

or jumps to 1 with probability

$$r(2,1) = P(min\{T_{R_1}, T_{R_2}\} > min\{T_{F_1}, T_{F_2}\}) = \frac{2/10}{2/4+2/10} = \frac{2}{7}$$

Suppose there are three machines working. The chain will remain there until $T_3 = min\{T_{R_1}, T_{F_1}, T_{F_2}, T_{F_3}\} \sim exp(\lambda_3 = \frac{1}{4} + \frac{3}{10})$. Then, the chain jumps to 4 with probability

$$r(3,4) = P(T_{R_1} < min\{T_{F_1}, T_{F_2}, T_{F_3}\}) = \frac{1/4}{1/4+3/10} = \frac{5}{11}$$

or jumps to 2 with probability

$$r(3,2) = P(T_{R_1} > min\{T_{F_1}, T_{F_2}, T_{F_3}\}) = \frac{3/10}{1/4+3/10} = \frac{6}{11}$$

Suppose there are four machines working. The chain will remain there until $T_4 = min\{T_{F_1}, T_{F_2}, T_{F_3}, T_{F_4}\} \sim exp(\lambda_4 = \frac{4}{10})$

Note that $\lambda_0 = \frac{1}{2}$, $\lambda_1 = \frac{3}{5}, \lambda_2 = \frac{7}{10}, \lambda_3 = \frac{11}{20}, \lambda_4 = \frac{2}{5}$. Hence, e.g. $q(1,2) = r(1,2)\lambda_1 = \frac{5}{6}\frac{3}{5} = \frac{1}{2}$.

Solve $\pi Q = 0$ gives the stationary distribution. The long-term expected number of working machines is $0\pi_0 + 1\pi_1 + 2\pi_2 + 3\pi_3 + 4\pi_4 = 2.76$. The desired probability for part 5 $= P_{42}(5) = (e^{5Q})_{42} = 0.188$.

Example 19. *During lunch hour, customers arrive at a fast-food restaurant at the rate of 120 customers per hour. The restaurant has one line, with three workers taking food orders at independent service stations. Each worker takes an exponentially distributed amount of time* *on average 1 minute to service a customer. Let X_t denote the number of customers in the restaurant (in line and being serviced) at time t. The process $\{X_t\}_{t\geq 0}$ is a continuous-time Markov Chain. Exhibit the generator matrix.*

Assume now that customers are turn away from the store if all three service stations are busy. Let Y_t denote the number of service stations busy at time t. Then, $\{Y_t\}_{t\geq 0}$ is a continuous-time Markov chain. Exhibit the generator matrix.

For the first question, note that the state space is $\mathbb{N}$. We start from 0 customer. Note that the average number of customer per minute is 2 so the time we stay in 0 customer is exponentially distributed with rate $\lambda_0 = 2$.

Suppose there is 1 customer. Then we either have 2 customers or we have 0 customer. First, we consider how long we will stay in 1 customer. Note that $T_1 = min\{T_{S_1}, T_{C_2}\} \sim exp(\lambda_1 = 1 + 2 = 3)$, where T_{S_1} = time to service customer 1, T_{C_2} = time for customer 2 come. Then, the chain jumps to 2 with probability

$$r(1,2) = P(T_{C_2} = T_1) = \tfrac{2}{1+2} = \tfrac{2}{3}$$

or jumps to 0 with probability

$$r(1,0) = P(T_{S_1} = T_1) = \tfrac{1}{1+2} = \tfrac{1}{3}$$

Note that $q(1,2) = r(1,2)\lambda_1 = 2$ and $q(1,0) = r(1,0)\lambda_1 = 1$

Suppose there are 2 customers. Let skip the detail and we have $\lambda_2 = min\{T_{S_1}, T_{S_2}, T_{C_3}\} = 1 + 1 + 2 = 4$

$$r(2,3) = P(T_{C_3} = T_2) = \tfrac{2}{4} = \tfrac{1}{2} \; r(2,1) = P(T_S = T_2) = \tfrac{1+1}{4} = \tfrac{1}{2}$$

Hence, we have $q(2,3) = r(2,3)\lambda_2 = 2$ and $q(2,1) = r(2,1)\lambda_2 = 2$

Suppose there are 3 customers. Similarly, $\lambda_3 = 1 + 1 + 1 + 2 = 5$ and we have

$$r(3,4) = P(T_{C_4} = T_2) = \tfrac{2}{5} \; r(3,2) = P(T_S = T_2) = \tfrac{1+1+1}{5} = \tfrac{3}{5}$$

Hence, we have $q(3,4) = r(3,4)\lambda_3 = 2$ and $q(2,1) = r(3,2)\lambda_3 = 3$. And forward and repeat the same process.

While for the second question, we restrict our state space to be $\{0,1,2,3\}$. Actually, we only need to modify if there are 3 customers as we cannot jump to 4 customers. Hence, $\lambda_3 = 1 + 1 + 1 = 3$ and $q(3,2) = 3$.

Example 20. *Let X_t be a standard Brownian motion on the real line. Answer the following questions.*

1. *What is the probability $P(X_3 \geq 1)$?*

2. *What is the probability $P(X_3 \geq X_1 + 1)$?*

3. *Find the mean and variance of $\int_0^1 X_t dt$. Hint: Assume that expectation and integration can be interchanged (which is indeed true). For the variance, note that $(\int_0^1 X_t dt)^2 = \int_0^1 \int_0^1 X_t X_s dt ds$.*

Note that $X_3 \sim N(0, 3)$ and hence $P(X_3 \geq 1) = 1 - \Phi(\frac{1}{\sqrt{3}}) = 0.282$.

Note that $P(X_3 \geq X_1 + 1) = P(X_3 - X_1 \geq 1) = P(X_2 \geq 1)$ by Markov property and $P(X_2 \geq 1) = 1 - \Phi(\frac{1}{\sqrt{2}}) = 0.240$.

Use the fact that $E[X_t] = 0$. Note that $E\left[\int_0^1 X_t dt\right] = \int_0^1 E[X_t]dt = \int_0^1 0dt = 0$. Moreover, recall that $Cov(X_s, X_t) = E(X_s X_t) = min\{s, t\}$. We have

$$\begin{aligned} Var\left(\int_0^1 X_t dt\right) &= E\left[\left(\int_0^1 X_t dt\right)^2\right] \\ &= E\left[\int_0^1 \int_0^1 X_t X_s dt ds\right] \\ &= \int_0^1 \int_0^1 E[X_t X_s] dt ds \\ &= \int_0^1 \int_0^1 min\{s, t\} dt ds \\ &= \frac{1}{3} \end{aligned}$$

Example 21. *Let X be a CTMC with infinitesimal generator:*

$$Q = \begin{bmatrix} -1 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -1 \end{bmatrix}$$

1. *Give an explicit formula for the transition matrix P_t by the eigendecomposition $Q = UDU^{-1}$.*

2. *Compute the stationary distribution π of the chain.*

3. *Find the transition probabilities for the embedded DTMC, and compute the stationary distribution for the embedded DTMC.*

For the first and second question, $P_t = e^{Qt} = U(e^{Dt})U^{-1}$ and the stationary distribution can be obtained from the first row of the U^{-1} matrix. So, It gives us $\pi = (1/3, 1/3, 1/3)$.

Recall that the routing probabilities $r(i,j)$ of the embedded Markov chain are $r(i,j) = q(i,j)/\lambda_j$. Then, the transition matrix of the embedded Markov chain is $\begin{bmatrix} 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \end{bmatrix}$.

Example 22. *Let $\{B_t\}_{t\geq 0}$ be a standard Brownian motion. Show that, for any $\lambda > 0$, $X_t = e^{\lambda B_t - \frac{1}{2}\lambda^2 t}$ satisfies the definition of martingale with respect to the Brownian motion.*

We just need to show X_t satisfy the definition of martingale, i.e. $E[X_t|B_u, u \leq s] = X_s$.

$$\begin{aligned} E[X_t|B_u, u \leq s] &= E[e^{\lambda B_t - \frac{1}{2}\lambda^2 t}|B_u, u \leq s] \\ &= e^{-\frac{\lambda^2 t}{2}t} E[e^{\lambda B_t}|B_u, u \leq s] \\ &= e^{-\frac{\lambda^2 t}{2}t} e^{\lambda B_s + \frac{1}{2}\lambda^2 (t-s)} \\ &= e^{\lambda B_s - \frac{1}{2}\lambda^2 s} \\ &= X_s \end{aligned}$$

While we use that $B_t|B_u, u \leq s \sim N(B_s, t-s)$ and $E[e^{\lambda N(\mu,v)}] = e^{\lambda\mu + \frac{\lambda^2}{2}v}$, it follows that $E[e^{\lambda B_t}|B_u, u \leq s] = e^{\lambda B_s + \frac{1}{2}\lambda^2(t-s)}$

Example 23. *Consider a standard Brownian motion started at $x = -3$.*

1. *Find the probability of reaching the level 2 before level -7.*
2. *When, on average, will the process leave the region between lines $y = 2$ and $y = -7$?*

By translation, the probability is equal to the probability that a standard Brownian motion started at 0 reaches y = 5 before y = -4 (a = 5, b = 4), and by the optional stopping theorem on quadratic martingale, we have $p = \frac{b}{a+b} = \frac{4}{9}$. Also, the expectation is $ab = 20$. Check the optional stopping theorem section for more details.

Example 24. *Let $\{Y_t\}_{t\geq 0}$ be a Brownian motion with drift $\mu \neq 0$ and variance σ^2 (i.e. $Y_t = \mu t + \sigma B_t$). Let $a, b > 0$.*

1. *Let $T = min\{t \geq 0 : Y_t = a$ or $Y_t = -b\}$. Argue that T is a stopping time; i.e. for any t, the occurrence of the event $\{T \leq t\}$ can be determined based only on the values of B up to time t.*
2. *Show that $X_t = e^{-\frac{2\mu Y_t}{\sigma^2}}$ is a martingale with respect to the Brownian motion B_t.*
3. *Use the previous parts as well as the optional stopping theorem to show that $E(e^{-\frac{2\mu Y_t}{\sigma^2}}) = 1$. Make sure to justify that T satisfies the conditions for the optional sampling theorem to hold.*

4. Use the previous part to find the probability that Y hits a before -b, for $a, b > 0$.

Note that $T = \{T_a, T_{-b}\}$, where T_a and T_{-b} are the first passage times of level a and $-b$, respectively. Then, for any $t \in (0, \infty)$, $\{T \leq t\} \subseteq \{T_a \leq t\} \cup \{T_{-b} \leq t\}$. But, both $\{T_a \leq t\}$ and $\{T_{-b} \leq t\}$ are events that can be determined from observing Y_s for $s \leq t$. Thus, we can determine if $\{T \leq t\}$ occurred or not from observing Y_s for $s \leq t$. In particular, since $Y_t = \mu t + \sigma B_t$, we can determine if $\{T \leq t\}$ occurred or not from observing B_s for $s \leq t$. This shows T is a stopping time.

Note that $Y_t = \mu t + \sigma B_t$. Let $X_t = e^{-\frac{2\mu Y_t}{\sigma^2}}$, which in terms of B_t takes the form:

$$X_t = e^{-\frac{2\mu Y_t}{\sigma^2}} = e^{\frac{2\mu}{\sigma} B_t - \frac{2\mu^2 t}{\sigma^2}}.$$

Note that X_t is also of the form $X_t = e^{\lambda B_t + \frac{\lambda^2}{2} t}$ with $\lambda = -2\mu/\sigma$. Therefore, X_t is a martingale.

By the optional stoppnig theorem,

$$E[X_T] = E[e^{-\frac{2\mu Y_T}{\sigma^2}}] = E[X_0] = 1$$

It remains to show that the optional stopping theorem can indeed be applied. So, we need to check that $E[X_t | 1_{t<T}] < \infty$. Indeed, this is the case because $|Y_t| \leq max\{a, b\}$, whenever $t < T$. Then, $X_t | 1_{t<T} \leq e^{2|\mu| max\{a,b\}/\sigma^2} < \infty$ and thus, $E[X_t | 1_{t<T}] < \infty$.

Since $Y_t = \begin{cases} a & \text{with prob } p \\ -b & \text{with prob } 1-p \end{cases}$, we have

$$E[e^{-\frac{2\mu Y_T}{\sigma^2}}] = pe^{-\frac{2\mu a}{\sigma^2}} + (1-p)e^{\frac{2\mu b}{\sigma^2}} = 1$$

We can then solve for p and obtain that

$$p = \frac{1 - e^{\frac{2\mu b}{\sigma^2}}}{e^{-\frac{2\mu a}{\sigma^2}} - e^{\frac{2\mu b}{\sigma^2}}}$$

Example 25. *Let $\{X_t\}_{t \geq 0}$ be a continuous-time Markov chain with state space 1,2 and rates q(1, 2) = 1 and q(2, 1) = 4. The following two parts must be done by "hand" (without using R or other software). (a) Find the transition probability matrix $P(t)$ and (b) the stationary distribution.*

Note that we have $Q = \begin{bmatrix} -1 & 1 \\ 4 & -4 \end{bmatrix}$. By some computation, it gives $\lambda = 0$ and $\lambda = 5$. With respect to different eigenvalues, we have eigenvectors [11] and $[1 - 4]$. It follows that $U = \begin{bmatrix} 1 & 1 \\ 1 & -4 \end{bmatrix}$ and we can find P_t by $Ue^{Qt}U^{-1}$.

To find the stationary distribution, we can take the limit as t goes to infinity or process as follows. For the stationary distribution we need solve the system $\pi Q = 0$. The first entry is

$$-\pi_1 + 4\pi_2 = 0$$

Adding the equation above to $\pi_1 + \pi_2 = 1$, we get

$$5\pi_2 = 1 \Rightarrow \pi_2 = \frac{1}{5}$$

Then, the stationary solution is $(\pi_1, \pi_2) = (4/5, 1/5)$.

Example 26. *A Markov chain $\{X_t\}_{t\geq 0}$ on {1, 2, 3, 4} has generator matrix:*

$$\begin{bmatrix} -2 & 1 & 1 & 0 \\ 1 & -3 & 1 & 1 \\ 2 & 2 & -4 & 0 \\ 1 & 2 & 3 & -6 \end{bmatrix}$$

Use technology as needed for the following:

1. *Find $P(X_1 = 3|X_0 = 1)$;*
2. *Find $P(X_5 = 1, X_4 = 4|X_{1.5} = 3, X_1 = 1)$*
3. *Find $P(X_{2.5} = 1)$ if the initial distribution of the Markov chain is $\alpha = [1/4, 1/4, 1/8, 3/8]$.*
4. *Find the long-term proportion of time the chain spends in state 1;*
5. *For the chain started in state 2, find the long-term probability that the chain is in state 3.*

Note that $P(X_1 = 3|X_0 = 1)$ and we have

$$\begin{aligned} P(X_5 = 1, X_4 = 4|X_{1.5} = 3, X_1 = 1) &= P(X_5 = 1, X_4 = 4|X_{1.5} = 3) \\ &= P(X_5 = 1|X_4 = 4, X_{1.5} = 3)P(X_4 = 4|X_{1.5} = 3) \\ &= P(X_5 = 1|X_4 = 4)P(X_4 = 4|X_{1.5} = 3) \\ &= P(X_1 = 1|X_0 = 4)P(X_{2.5} = 4|X_0 = 3) \\ &= P_{4,1}(1)P_{3,4}(2.5) \end{aligned}$$

For part 3, we need to find $[\alpha P_{2.5}]_1 = 0.407$.

For part 4, this is the first entry of the limiting distribution of the chain, and for part 5, it is the limit of $P_{2,3}(t)$.