Recap: Brownian Motion

- $\{X(t),\ t \geq 0\}$ is called a Brownian motion if
  1. $X(0) = 0$
  2. The process has stationary and independent increments
  3. For every $t > 0$, $X(t)$ is Gaussian with mean 0 and variance $\sigma^2 t$
- If $\sigma^2 = 1$, it is called standard Brownian Motion
- $Y(t) = X(t) + \mu$ is called Brownian motion with a drift

Recap: Some properties

- An important result is that Brownian motion paths are continuous
- The autocorrelation and autocovariance of the process is

$$R(t_1, t_2) = \sigma^2 \min(t_1, t_2) \ = \mathsf{Cov}(X(t_1), X(t_2))$$

- All n^{th} order distributions are Gaussian: $X(t_1), \cdots, X(t_n)$ are jointly Gaussian.

Recap: Gaussian Processes

- A continuous-time continuous-state process $\{X(t),\ t \geq 0\}$ is said to be a Gaussian process if for all n and all $t_1, t_2, \cdots, t_n$, we have that $X(t_1), \cdots, X(t_n)$ are jointly Gaussian.
- A general Gaussian process is specified by the mean function and the variance and covariance functions
- The Brownian motion is an example of a Gaussian Process
- The Brwonian motion is a Gaussian process with

$$E[X(t)] = 0, \quad \mathsf{Var}(X(t)) = \sigma^2 t, \quad \mathsf{Cov}(X(s), X(t)) = \sigma^2 \min(s,t)$$

Recap: Conditional Densities

- For $s < t$, $f_{X(s)|X(t)}$ is Gaussian and

$$E[X(s)|X(t)] = \frac{s}{t}\, X(t); \;\; \mathsf{Var}(X(s)|X(t)) = \frac{s}{t}\, (t-s)$$

- For $s < t$, we also showed $f_{X(t)|X(s)}$ is Gaussian with

$$E[X(t) \mid X(s)] = X(s); \;\; \mathsf{Var}(X(t)|X(s)) = (t-s)$$

Recap: Hitting Times

- Let T_a denote the first time Brownian motion hits a.

$$Pr[T_a \leq t] = \frac{2}{\sqrt{2\pi}} \int_{|a|/\sqrt{t}}^{\infty} e^{-\frac{y^2}{2}} \, dy$$

- By continuity of sample paths,

$$Pr[\max_{0 \leq s \leq t} X(s) \geq a] = Pr[T_a \leq t] = 2Pr[X(t) \geq a]$$

Recap: Geometric Brownian Motion

- Let $\{Y(t),\ t \geq 0\}$ be a Brownian motion with drift. Define
$$X(t) = e^{Y(t)}$$
- Then, $\{X(t),\ t \geq 0\}$ is called geometric Brownian motion. It is useful in mathematial finance

- We have considered three random processes
- (discrete-time) Markov Chain
  – Example of Discrete-time discrete-state process
- Continuous time Markov Chains (e.g., Poisson Process)
  – Example of continuous-time discrete-state process
- Brownian Motion
  – Example of continuous-time continuous-state process
- We need an example of discrete-time continuous-state process!
- Any sequence of continuous random variables would be a discrete-time continuous-state process

- In general, any 'stochastic' algorithm would generate discrete-state continuous time process.
- If an algorithm uses a random step, then the algorithm would be like

$$X(n+1) = X(n) + \eta_n \, G(X(n), \xi(n))$$

  where $\xi(n)$ would be some random variable which may be dependent on $X(n)$.
- Many algorithms can be written in this general form.
- The $X(n), n = 0, 1, \cdots$ would be a discrete-time continuous-state stocahstic process
- We consider an important class of such processes

- Let $\{X_n,\ n = 0, 1, \cdots\}$ be a discrete-time continuous-state process.
- It is called a **martingale** if $E|X_n| < \infty,\ \forall n$ and

$$E[X_{n+1} \mid X_n, \cdots, X_0] = X_n,\ \forall n$$

- One can think of Martingale as a general fair gambling game.
- When X_n is a martingale, we have

$$E[X_{n+1}] = E[X_n],\ \forall n$$

Example

- Sequence of sums of independent random variables forms a Martingale
- Suppose Z_i are iid with $Pr[Z_i = +1] = Pr[Z_i = -1] = 0.5$. Let

$$X_n = \sum_{i=1}^{n} Z_i \quad \Rightarrow \quad X_{n+1} = X_n + Z_{n+1}$$

- Since $EZ_i = 0, \ \forall i$,

$$E[X_{n+1} \mid X_n, \cdots] = E[X_n + Z_{n+1} | X_n \cdots] = X_n + E[Z_{n+1} | X_n \cdots] = X_n$$

- Hence, X_n is a martingale.

Example

- Suppose $X_1, X_2, \cdots$ are iid with $E[X_i] = 1$.
- Let $Z_n = \prod_{i=1}^n X_i$.
- Since $Z_{n+1} = Z_n \; X_{n+1}$,

$$\begin{aligned} E[Z_{n+1} | Z_n, \cdots, Z_1] &= E[Z_n X_{n+1} | Z_n, \cdots, Z_1] \\ &= Z_n \; E[X_{n+1} | Z_n, \cdots, Z_1] \\ &= Z_n, \quad \text{since} \quad X_i iid, E[X_i] = 1 \end{aligned}$$

- Thus, Z_n is a martingale

- Given $\{X_n,\ n = 1, 2, \cdots\}$ and $E|X_n| < \infty,\ \forall n$, It is called a martingale if

$$E[X_{n+1} \mid X_n, \cdots, X_1] = X_n,\ \forall n$$

- We may often actually show

$$E[X_{n+1} \mid Y_n, \cdots, Y_1] = X_n,\ \forall n$$

  Where $Y_1, \cdots, Y_n$ determine $X_1, \cdots, X_n$ for all n.
- This is justified by the following two identities

$$E[\ E[X_{n+1} \mid X_n, \cdots, X_1,\ Y_n, \cdots, Y_1] \mid X_1, \cdots, X_n] \\ = E[X_{n+1} \mid X_n, \cdots, X_1]$$

for any X_i, Y_i by properties of conditional expectation

$$E[X_{n+1} \mid X_n, \cdots, X_1,\ Y_n, \cdots, Y_1] = E[X_{n+1} \mid Y_n, \cdots, Y_1]$$

when $Y_1, \cdots, Y_n$ determine $X_1, \cdots, X_n$ for all n

A property of Conditional Expectation

- Conditional expectation satisfies

$$E[\ E[h(X)|Y,Z]\mid Y] = E[h(X)|Y]$$

  Note that all these can be random vectors.
- Let

$$\begin{aligned} g_1(Y,Z) &= E[h(X)|Y,Z] \\ g_2(Y) &= E[g_1(Y,Z)|Y] \end{aligned}$$

  We want to show $g_2(Y) = E[h(X)|Y]$

- Recall: $g_1(Y,Z) = E[h(X)|Y,Z]$, $g_2(Y) = E[g_1(Y,Z)|Y]$

$$\begin{aligned}
g_2(y) &= \int g_1(y,z)\ f_{Z|Y}(z|y)\ dz \\
&= \int \left[\int h(x)\ f_{X|YZ}(x|y,z)\ dx \right] f_{Z|Y}(z|y)\ dz \\
&= \int h(x)\ \left[\int f_{X|YZ}(x|y,z)\ f_{Z|Y}(z|y)\ dz \right]\ dx \\
&= \int h(x)\ \left[\int f_{XZ|Y}(x,z|y)\ dz \right]\ dx \\
&= \int h(x)\ f_{X|Y}(x|y)\ dx
\end{aligned}$$

- Thus we get

$$E[\ E[h(X)|Y,Z]\ |\ Y] = E[h(X)|Y]$$

Example

- We saw that partial sums of iid X_i with $E[X_i] = 0$ is a martingale.
- We can generalize this.
- Let $X_1, X_2, \cdots$ any sequence of random variables.
- Let $W_i = X_i - E[X_i \mid X_1, \cdots, X_{i-1}]$.
- Now $E[W_i] = 0, \forall i$ though W_i may not be independent.
- Let $Z_n = \sum_{i=1}^n W_i$.
- We can show that $Z_n,\ n = 1, 2, \cdots$ is a martingale, assuming that $E[|Z_n|] < \infty$.

- We have
$Z_n = \sum_{i=1}^{n} W_i = \sum_{i=1}^{n} X_i - E[X_i \mid X_1, \cdots, X_{i-1}]$
- Note that $X_1, \cdots, X_n$ determine $Z_1, \cdots, Z_n, \ \forall n$
- We have, $Z_{n+1} = Z_n + X_{n+1} - E[X_{n+1}|X_1, \cdots, X_n]$.

$$\begin{aligned}
E[Z_{n+1}|X_1, \cdots, X_n] &= E[Z_n + W_{n+1}|X_1, \cdots, X_n] \\
&= Z_n + E[X_{n+1}|X_1, \cdots, X_n] - \\
&\quad E[X_{n+1}|X_1, \cdots, X_n] \\
&= Z_n
\end{aligned}$$

- This shows that Z_n is a martingale

- Let N be a positive integer valued random variable with $P[N < \infty] = 1$.
- N is said to be a stopping time for the process $Z_1, Z_2, \cdots$ if

  the event $[N = n]$ is determined by the random variable $Z_1, \cdots, Z_n$.
- If we know the values of $Z_1, \cdots, Z_n$, then we can say whether or not $N = n$.
- The idea is that we can decide to stop the process at N and the decision to stop cannot anticipate the future

- Let N be a stopping time for the process $\{Z_n,\ n \geq 1\}$.
- Define
$$\bar{Z}_n = \begin{cases} Z_n & \text{if } n \leq N \\ Z_N & \text{if } n > N \end{cases}$$
- We call $\{\bar{Z}_n,\ n \geq 1\}$ the stopped process

- Let $\{Z_n,\ n \geq 1\}$ be a martingale and let N a stopping time for it.
- **Theorem**: The stopped Process is a martingale
  That is, a stopped martingale is a martingale
- Note that, by definition, $\bar{Z}_1 = Z_1$.
- So, if the stopped process is a martingale, then,

$$E[\bar{Z}_n] = E[\bar{Z}_1] = E[Z_1], \forall n$$

- Recall that for a martingale, $E[Z_n] = E[Z_1],\ \forall n$
- The theorem essentially says there is no strategy to have a positive expectation from a fair gambling game.

- $\{X_n,\ n = 0, 1, \cdots\}$ and $E|X_n| < \infty,\ \forall n$ is called a martingale if

$$E[X_{n+1} \mid X_n, \cdots, X_0] = X_n,\ \forall n$$

- It is called a supermartingale if

$$E[X_{n+1} \mid X_n, \cdots, X_0] \leq X_n,\ \forall n$$

- It is called a submartingale if

$$E[X_{n+1} \mid X_n, \cdots, X_0] \geq X_n,\ \forall n$$

- Note that $E[X_{n+1}] \leq E[X_n]$ for supermartingales and $E[X_{n+1}] \geq E[X_n]$ for submartingales.
- In the above, the conditioning random variables can be another sequence Y_i if $Y_1, \cdots, Y_n$ determine $X_1, \cdots, X_n$

- A useful result is the martingale convergence theorem.

  Martingale convergence theorem: If X_n is a martingale with $\sup_n E|X_n| < \infty$ then X_n converges almost surely to a rv X which will have finite expectation.
  A positive supermartingale also converges almost surely
- If X_n are bounded, then the condition is always true and the almost sure convergence implies convergence in the mean.
- This is often useful in dealing with many sequences of random variables such as a stochastic algorithm.

Example

- We consider a simple algorithm for a two armed bandit problem.
- 2 arms. Response is binary (1 for reward).
- d_i - prob of reward for arm-i. Do not know d_i
- need to play and find which is better arm.
- We choose arm-1 with prob $p(k)$ (and hence arm-2 with prob $(1 - p(k))$ at iteration k and update $p(k)$ based on the outcome.

- $p(k)$ denotes probability of choosing arm-1 at k and $b(k)$ denotes the response.
- Our algorithm for the 2-armed bandit problem is

$$\begin{aligned} p(k+1) &= p(k) + \lambda(1 - p(k)) \text{ if arm 1 chosen, } b(k) = 1 \\ &= p(k) - \lambda p(k) \text{ if arm 2 is chosen and } b(k) = 1 \\ &= p(k) \text{ if } b(k) = 0 \end{aligned}$$

- We want to know whether the algorithm converges.

- The algorithm is

$$
\begin{aligned}
p(k+1) &= p(k) + \lambda(1 - p(k)) \;\; \text{if arm 1 chosen, } b(k) = 1 \\
&= p(k) - \lambda p(k) \;\; \text{if arm 2 is chosen and } b(k) = 1 \\
&= p(k) \;\; \text{if } b(k) = 0
\end{aligned}
$$

- We get

$$
\begin{aligned}
& E[p(k+1) - p(k) | p(k)] \\
& \quad = \lambda(1 - p(k)) \; Pr[b(k) = 1, \text{arm 1} \mid p(k)] \\
& \qquad - \lambda p(k) \; Pr[b(k) = 1, \text{arm 2} \mid p(k)] \\
& \quad = \lambda(1 - p(k)) \; Pr[b(k) = 1 \mid \text{arm 1}, p(k)] \; Pr[\text{arm 1} \mid p(k)] \\
& \qquad - \lambda p(k) \; Pr[b(k) = 1 \mid \text{arm 2}, p(k)] \; Pr[\text{arm 2} \mid p(k)]
\end{aligned}
$$

- This gives us

$$
\begin{aligned}
E[p(k+1) - p(k)|p(k)] &= \lambda(1 - p(k))\ d_1\ p(k) \\
&\qquad -\lambda p(k)\ d_2\ (1 - p(k)) \\
&= \lambda\, p(k)(1 - p(k))\ (d_1 - d_2) \\
&\geq 0, \quad \text{if} \quad d_1 > d_2
\end{aligned}
$$

$$
\Rightarrow \quad E[p(k{+}1) \mid p(k)] \geq p(k) \quad \Rightarrow E[p(k{+}1)] \geq E[p(k)],\ \forall k
$$

- This also shows $p(k)$ is a submartingale.
- Here, $p(k)$ is bounded and $1 - p(k)$ is a supermartingale.
- So, we can conclude, the algorithm converges almost surely

- We considered martingales as an example of discrete-time continuous-state processes
- Stochastic iterative algorithms generate such processes.
- Martingales can be useful in analyzing convergence of many stochastic algorithms
- While we mentioned only discrete-time martingales, one can similarly have continuous-time martingales that satisfy

$$E[X(t)|(X(s'), 0 \leq s' \leq s < t] = X(s)$$