Ch. 20 Brownian motion and Itô’s lemma

Let \((\Omega, \mathcal{F}, P)\) be a probability space.

A stochastic process on \(\Omega\) is a set of functions (random variables) \(X_t: \Omega \to \mathbb{R}\) for \(t \in S\), where often \(S = \mathbb{N}\) (a discrete process) or \(S = [0, \infty)\) (a continuous process).

It is often useful to have a filtration, i.e., a set of \(\sigma\)-algebras \(\mathcal{F}_t\) on \(\Omega\) such that if \(s, t \in S\) and \(s \leq t\), then \(\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}_t\), and to require that the process is adapted to this filtration in the sense that for all \(b \in \mathbb{R}\) and \(t \in S\), we have

\[ X_t^{-1}((\infty, b)) = \{ \omega \in \Omega : X_t(\omega) < b \} \in \mathcal{F}_t. \]

We often write \(X(t)\) instead of \(X_t\).

Such a process is called a Brownian motion (1-dimensional) if \(S = [0, \infty)\) and

- \(X_0 = 0\) with probability 1 (a.s. = almost surely)
- if \(t, s, s_2 > 0\) then the increment \(X(t+s) - X(t)\) is independent of \(X(t) - X(t-s_2)\)
- if \(t, s > 0\) then \(X(t+s) - X(t) \sim \mathcal{N}(0, s)\)
- with probability 1, the map \(\omega \to X_t(\omega)\) is a continuous function of \(t\) on \([0, \infty)\).
Note that we did not specify what \( \Omega, F, F_t \) are. They can be anything as long as the above properties are satisfied.

On the one hand, it is difficult to prove that any Brownian motion exists, but this has been done.

The space \( \Omega \) must be large. An often used example is \( C([0, \infty], \mathbb{R}) \), the set of all continuous real-valued functions on \([0, \infty)\). The measure \( P \) is then the so-called Wiener measure. Brownian motion is also called the Wiener process (after Norbert Wiener).

One can show that there are many Brownian motions, and that when we start with many sequences of random variables and average them in a suitable way, we get a Brownian motion as the limit.

We may denote the Brownian motion by many different letters, not necessarily always by \( X \) or \( X_t \) or \( X(t) \).
Let us consider a Brownian motion $Z(t)$. Then for $h > 0$,

$$Z(t+h) - Z(t) = \sqrt{h} Y(t)$$

where

$$Y(t) \sim \mathcal{N}(0,1).$$

If $h$ is small, we adopt the notation

$$h = dt, \quad Z(t+h) - Z(t) = dZ(t).$$

Then this reads

$$dZ(t) = Y(t) \sqrt{dt}.$$

Performing formal computations with this notation is useful. We may also write

$$Z(T) = Z(0) + \int_0^T dZ(t).$$

The quadratic variation of $Z(T)$ is

$$h = T/n,$$

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \left( Z(ih) - Z((i-1)h) \right)^2 = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} Y_{ih}^2 = \sqrt{h} Y_{ih}$$

Instead of trying to evaluate this sum in general, we take the heuristic view (which can be justified) that the $Y_{ih}$ are i.i.d. each with probability $1/2$. Then $Y_{ih}^2 = 1$, so the sum is

$$nh = T.$$

Thus the quadratic variation of $Z(T)$ is $T$. Just for the purpose of this calculation...
An arithmetic Brownian motion is a process of the form:

\[ X(t+h) - X(t) \sim \alpha h + \sigma Y(t+h) \sqrt{h}, \]

where \( Y(t+h) \sim W(0,1) \).

We write:

\[ dX(t) = \alpha \, dt + \sigma \, dZ(t) \]

where \( Z(t) \) is a Brownian motion. We have

\[ X(T) - X(0) \sim W(\alpha T, \sigma^2 T). \]

Also:

\[ X(T) = X(0) + \int_0^T \alpha \, dt + \int_0^T \sigma \, dZ(t). \]

The Ornstein–Uhlenbeck process

\[ dX(t) = \lambda (\alpha - X(t)) \, dt + \sigma \, dZ(t) \]

can be used to describe a process that tries to revert to its mean.
A geometric Brownian motion is a process

\[ dX(t) = \alpha X(t) \, dt + \sigma X(t) \, dZ(t), \]

i.e.,

\[ \log X(t) = \frac{dX(t)}{X(t)} = \alpha \, dt + \sigma \, dZ(t), \]

\[ X(T) - X(0) = \int_0^T \alpha X(t) \, dt + \int_0^T \sigma X(t) \, dZ(t). \]

Then

\[ \ln X(t) \sim N \left( \ln X(0) + (\alpha - \frac{1}{2} \sigma^2) t, \sigma^2 t \right), \]

\[ X(t) = X(0) \, e^{(\alpha - \frac{1}{2} \sigma^2) t + \sigma \sqrt{t} Z} \]

where \( Z \sim N(0,1) \).

Thus geometric Brownian motion gives rise to a log-normal distribution.
Consider the two terms on the right hand side of the equation for the arithmetic Brownian motion:

\[ X(t+h) - X(t) = \alpha h + \sigma \sqrt{h} Y(t) \]

We have

\[ \frac{\alpha h}{\sigma \sqrt{h}} = \frac{\alpha}{\sigma} \sqrt{h} \]

which is small when \( h \) is small. So when \( h \) is small, the random term \( \sigma \sqrt{h} Y(t) \) dominates. If \( h \) is large (say, \( h \geq 1 \) year), then the drift term \( \alpha h \) dominates (unless \( \alpha = 0 \)).

When \( h = dt \) is small, \( \sqrt{dt} \) is much larger than \( dt \). To get an idea of the relative magnitude of terms in the above and in other similar expressions, it is useful to think of each term having a factor of the form \( (dt)^{\frac{1}{2}} \) for some exponent \( \gamma \), and to take into account that the smaller \( \gamma \) is, the larger the term is (when \( dt \) is assumed to be very small). In this way we see, e.g., that (1) implies that

\[ (X(t+h)-X(t))^2 \approx \sigma^2 h Y^2 \approx \sigma^2 \]

if we assume \( Y = \pm 1 \), i.e., \( Y^2 = 1 \).
People sometimes use multiplication rules like

\[ \frac{dZ}{dt} = 0 \]
\[ (dZ)^2 = dt \]

The idea being to get \((dZ)^2 = 0\) if \(r > 1\), such quantities being too small to count as \(dt \to 0\).

**Modeling asset prices**

In accordance with the connection mentioned above, between Brownian motion and the lognormal distribution, we can model the price of one asset such as a stock by

\[ \frac{dX(t)}{X(t)} = (\mu - \delta) \, dt + \sigma \, dZ(t) \]

where \(Z(t)\) is a Brownian motion.

Suppose now that we have \(n\) correlated assets, and we model their prices by

\[ \frac{dX_i}{X_i} = (\mu_i - \delta_i) \, dt + \sigma_i \, dZ_i, \quad 1 \leq i \leq n \]

where each \(Z_i\) is a Brownian motion.
To construct $\mathbf{Z}$ so that we get prescribed correlations, we assume it to be known that it is possible to construct $n$ pairwise independent Brownian motions $W_i$, $1 \leq i \leq n$, so that (again a multiplication rule)

$$dW_i \cdot dW_j = 0 \quad \text{whenever} \quad i \neq j \quad (1 \leq i, j \leq n).$$

Using numbers $\lambda_{ij}$ to be determined later, define

$$dZ_i(t) = \sum_{k=1}^{n} \lambda_{ik} dW_k(t).$$

We have

$$\text{Var} \left[ dZ_i(t) \right] = \text{Var} \left[ \sum_{k=1}^{n} \lambda_{ik} dW_k(t) \right]$$

since the $W_k$ are independent, i.e., their correlations are zero as desired, if $\forall k$ we have $\sum_{k=1}^{n} \lambda_{ik}^2 = 1$.

The mean of $dZ_i$ is $0$ since the mean of each $dW_k$ is zero.

To determine the correlation between $dZ_i$ and $dZ_j$, we calculate

$$dZ_i \cdot dZ_j = \sum_{k=1}^{n} \lambda_{ik} dW_k \sum_{l=1}^{n} \lambda_{lj} dW_l = \sum_{k=1}^{n} \lambda_{ik} \lambda_{jk} \left( \frac{dW_k}{dt} \right)^2 = \rho_{ij} dt,$$

where $\rho_{ij} = \sum_{k=1}^{n} \lambda_{ik} \lambda_{jk}$. Here we used $dW_k \cdot dW_l = 0$ if $k \neq l$. 
If we write $L$ and $R$ for the $n$-by-$n$ matrices with elements $L_{ij}$ and $R_{ij}$, we see that

$$R = L L^T.$$  

The matrix $R$ must be positive definite.

If $L$ is required to be lower triangular, then if $R$ is given, there is a unique $L$ with $R = L L^T$ (Cholesky decomposition from linear algebra). "Lower triangular" means that

$$dZ_1 = dW_1,$$

$$dZ_2 = dW_1 + dW_2,$$

$$dZ_3 = dW_1 + dW_2 + dW_3,$$

and so on.

This is a lot simpler than general linear combinations.

If now

$$dX_i / X_i = (x_i - \delta_i) \, dt + \Theta_i \, dZ_i,$$

then

$$dX_i^2 = X_i^2 \left( (x_i - \delta_i)^2 \, dt + \Theta_i^2 \, dZ_i^2 + 2 (x_i - \delta_i) \Theta_i \, dt \, dZ_i \right),$$

$$= X_i(t) \, \delta_i^2 \, dt,$$

discarding terms that involve $(dt)^2$ for $R > 1$.

Note the multiplication rules we mentioned before.
Similarly, if \( i \neq j \),

\[
dx_i \, dx_j = x_i x_j \left[ (x_i - \delta_i) (x_j - \delta_j) \, dt^2 + \delta_{ij} \, d\bar{z}_i \, d\bar{z}_j \right] + (x_i - \delta_i) \delta_j \, dt \, d\bar{z}_j + (x_j - \delta_j) \delta_i \, dt \, d\bar{z}_i
\]

\[
= x_i x_j \delta_{ij} \, \delta_{ij} \, dt
\]

since \( d\bar{z}_i \, d\bar{z}_j = \delta_{ij} \, dt \). The higher order terms in \( dt \) are omitted.
Itô's Lemma

Suppose that \( f(t, x) \) is a twice continuously differentiable function of \( t \) and \( x \).

We want to substitute for \( x \) a variable \( X \) for which

\[
dx = \alpha \ dt + \sigma \ dZ
\]

for a Brownian motion \( Z \). We want to find \( df(t, X) \) so that we keep terms that are \( dt \) or larger, and discard smaller terms.

The ordinary Taylor expansion of \( f \) gives

\[
df = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dx + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} dx^2 + \ldots
\]

We do not need second order terms involving \( dx dt \) or \( dt^2 \) since these would be smaller than \( dt \). To this, we substitute

\[
dx = \alpha \ dt + \sigma \ dZ
\] and

\[
(dx)^2 = \alpha^2 \ dt^2 + 2 \alpha \sigma \ dt \ dZ + \sigma^2 \ dZ^2 = \sigma^2 \ dZ^2
\]

since the other terms are smaller than \( dt \). Also \( \sigma^2 \ dZ^2 = \sigma^2 \ dt \).
Thus we get
\[ df = \frac{dt}{dt} + \frac{df}{dx} \left( \alpha dt + \sigma dZ \right) + \frac{1}{2} \frac{\sigma^2}{dx^2} dt \]
\[ = \left( \frac{2f}{dx} + \alpha \frac{f}{dx} + \frac{\sigma^2}{2} \frac{d^2 f}{dx^2} \right) dt + \sigma \frac{df}{dx} dZ. \]

This result is known as Itô's lemma after Kiyoshi Itô.

Suppose now that
\[ df(x) = \frac{dx}{X} = (x - \delta) dt + \sigma dZ \]

We apply Itô's lemma to \( df(x) \) instead of \( f \) and obtain (going through a calculation similar to the one above)
\[ dX = (x - \delta) X dt + \sigma X dZ, \]
\[ (dX)^2 = \sigma^2 X^2 dt, \]

hence
\[ df = \left( \frac{df}{dt} + (x - \delta) \frac{df}{dx} \frac{d^2 f}{dx^2} + \frac{\sigma^2 X^2}{2} \frac{d^2 f}{dx^2} \right) dt + \sigma X \frac{df}{dx} dZ. \]
If we take \( d \) to be the price of a European call option, \( C = C(S,t) \), denote \( X \) by \( S \) (stock price), so that

\[
\frac{dS}{S} = (\alpha - \delta) dt + \sigma dZ
\]

and if we use subscripts to denote partial derivatives, so that

\[
\begin{align*}
\frac{\partial C}{\partial t} &= \psi, \\
\frac{\partial C}{\partial S} &= \sigma, \\
\frac{\partial^2 C}{\partial S^2} &= \gamma,
\end{align*}
\]

we can write the consequence of Itô’s lemma in this case in the form

\[
dC(S,t) = \left( \frac{\partial C}{\partial t} + (\alpha - \delta) S \frac{\partial C}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} \right) dt + \sigma S \frac{\partial C}{\partial S} dZ.
\]

*Note that certain option Greeks are equal to*

\[
\begin{align*}
\Delta &= \psi, \\
\Gamma &= \gamma, \\
\Theta &= \psi
\end{align*}
\]

so that we get the approximation

\[
dC = \left( \theta + (\alpha - \delta) S \Delta + \frac{1}{2} \sigma^2 S^2 \Gamma \right) dt + \sigma S \Delta dZ.
\]
Multivariate Itô’s lemma

Suppose that the price of an option (or other financial instrument), say \( C \), depends on asset prices \( S_1, \ldots, S_n \) that are correlated, and on time \( t \). We assume that

\[
\frac{dS_i(t)}{S_i} = \alpha_i dt + \sigma_i dZ_i, \quad 1 \leq i \leq n
\]

(i.e., not \( \alpha_i = \delta_i \) for shorter notation).

Suppose

\[
dZ_i \cdot dZ_j = \rho_{ij} dt.
\]

Let \( C \) be a twice continuously differentiable function in all of its variables. By the ordinary Taylor expansion of \( C \) we get

\[
dC = \sum_{i=1}^{n} \frac{\partial C}{\partial S_i} dS_i + \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 C}{\partial S_i \partial S_j} dS_i dS_j + C_t dt.
\]

Next,

\[
dS_i dS_j = S_i S_j (\alpha_i dt + \sigma_i dZ_i)(\alpha_j dt + \sigma_j dZ_j)
\]

\[
= S_i S_j \sigma_i \sigma_j \rho_{ij} dt.
\]

Here

\[
dC = \left( \sum_{i=1}^{n} \frac{\partial C}{\partial S_i} \alpha_i S_i + C_t + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 C}{\partial S_i \partial S_j} S_i S_j \sigma_i \sigma_j \rho_{ij} \right) dt + \sum_{i=1}^{n} \frac{\partial C}{\partial S_i} \sigma_i S_i \sigma_i dZ_i.
\]
Since $E(dZ_i) = 0$, the expected change in $C$ in time $dt$ is

$$\frac{dt}{\partial C} = \sum_{i=1}^{n} \alpha_i S_i C_i + \frac{1}{2} \sum_{i,j=1}^{n} \sigma_{ij} S_i S_j C_i C_j.$$

The Sharpe ratio of an asset is $\frac{\alpha - r}{\sigma}$, where $\alpha$ is the rate of expected return, $r$ is the risk-free interest rate, and $\sigma$ is the volatility of the asset price.

Suppose that two assets depend, as far as the random component of their prices are concerned, on the same Brownian motion $Z$. Then,

$$dS_1 = \alpha_1 S_1 dt + \sigma_1 S_1 dZ,$$

$$dS_2 = \alpha_2 S_2 dt + \sigma_2 S_2 dZ.$$

Construct a portfolio where we buy $\frac{1}{\sigma_2 S_2}$ shares of asset 1, short $\frac{1}{\sigma_1 S_1}$ shares of asset 2, and invest (or borrow depending on the sign) the price difference $\frac{1 - 1}{\sigma_1 \sigma_2}$ at the risk-free rate.
The value of the total portfolio is

\[
\frac{1}{\sigma_1 S_1} dS_1 + \frac{1}{\sigma_2 S_2} dS_2 + \left( \frac{1 - \frac{1}{\sigma_1}}{\sigma_2} \right) r dt
\]

\[
= \frac{\alpha_1}{\sigma_1} dt + dZ = \left( \frac{\alpha_2}{\sigma_2} - \frac{1}{\sigma_2} \right) dt + \left( \frac{1 - \frac{1}{\sigma_1}}{\sigma_2} \right) r dt
\]

\[
= \left( \frac{\alpha_1 - r}{\sigma_1} - \frac{\alpha_2 - r}{\sigma_2} \right) dt.
\]

If \( \frac{\alpha_1 - r}{\sigma_1} \) \( > \) \( \frac{\alpha_2 - r}{\sigma_2} \), we can make a risk-free profit by this portfolio. If \( \frac{\alpha_1 - r}{\sigma_1} < \frac{\alpha_2 - r}{\sigma_2} \), we can make a risk-free profit by doing the opposite. To avoid arbitrage, we must have

\[
\frac{\alpha_1 - r}{\sigma_1} = \frac{\alpha_2 - r}{\sigma_2}
\]

i.e., two asset depending on the same \( dZ \) (such as a stock and an option based on the stock) must have the same Sharpe ratio (ratio of the risk premium \( \alpha - r \) to volatility \( \sigma \) ).