1 Stochastic Differential Equations and the Black-Scholes PDE

— Derivation of the BS PDE via Ito calculus

Motivations:

1. Ito calculus provides us an elegant tool to deal with many financial derivative problems, e.g., the meaning of hedging in continuous time can be easily clarified

2. PDE methods lead to efficient computational methods analytically and numerically, e.g., trinomial tree method is merely a numerical method for solving these PDE’s

1.1 Brownian Motion

Binomial models gives rise to a lognormal stock price dynamics, i.e.,

\[ \log \frac{S(t_2)}{S(t_1)} \text{ is a Gaussian random variable with} \]

\[ \text{mean } \mu (t_2 - t_1) \text{ and variance } \sigma^2 (t_2 - t_1) \]

Consider the following limit process:

As \( \Delta t \to 0 \), \( x_n \to x(t) \)

then \( x(t) \) is a stochastic process, i.e., a time-dependent random variable.
1.1.1 Brownian Motion $W(t)$:

Brownian motion $W(t)$:

1. $W(t_2) - W(t_1)$ is a Gaussian random variable with mean 0 and variance $t_2 - t_1$.

2. Distinct intervals give rise to independent random variables, i.e.,

$$W(t_2) - W(t_1) \text{ and } W(t_4) - W(t_3)$$

are statistically independent if the intervals $[t_1, t_2]$ and $[t_3, t_4]$ do not overlap.

3. $W(0) = 0$

Note that

1. These properties determine a unique stochastic process, i.e.,

Given any $t_1, t_2, \ldots, t_n \forall n$, then

$$P(W(t_1), W(t_2), \ldots, W(t_n))$$

$$P(W(t_i), W(t_j), \ldots \forall i, j)$$

$$P(W(t_i), W(t_j), W(t_k), \ldots \forall i, j, k)$$

$$\cdots$$

$$P(W(t_1), W(t_2), \ldots W(t_n))$$

are specified.

2. Any realization:

$$t \mapsto W(t)$$

is continuous but nowhere differentiable since

$$E \left| \frac{\Delta W}{\Delta t} \right| \sim \frac{\Delta t^{1/2}}{\Delta t} \sim \frac{1}{\Delta t^{1/2}}$$

as $\Delta t \to 0$.
3. Define
\[
\Delta W = W(t_2) - W(t_1) \\
\Delta t = t_2 - t_1 \\
\mathbb{E} [\|\Delta W^n\|] = C_n |\Delta t|^{n/2}, \quad n = 1, 2, \ldots
\]
and
\[
\mathbb{E} [\Delta W] = 0 \\
\mathbb{E} [(\Delta W)^2] = \Delta t \\
Var [\Delta W] = \Delta t \\
Var [(\Delta W)^2] = 2 (\Delta t)^2
\]
the last of which can be seen as follows:
\[
Var [(\Delta W)^2] = \mathbb{E} [(\Delta W)^4] - \mathbb{E} [(\Delta W)^2]^2 \\
= 3 (\Delta t)^2 - (\Delta t)^2 \\
= 2 (\Delta t)^2
\]
Note that $(\Delta W)^2$ has mean $\Delta t$ and variance $(\Delta t)^2$. Since this variance/mean $\rightarrow 0$ as $\Delta t \rightarrow 0$, $(\Delta W)^2$ looks more and more "deterministic" as $\Delta t \rightarrow 0$. In the limit, we can express this fact by
\[
(dW(t))^2 = dt
\]
which has a strong ramification for the derivation of Ito formula later.

4. Stock price dynamis can be written as
\[
\log \frac{S(t)}{S(0)} = \mu t + \sigma W(t)
\]
\[
\Rightarrow 
S(t) = S(0) e^{\mu t + \sigma W(t)}
\]

1.2 Stochastic Differential Equations and Ito’s Lemma

1.2.1 Deterministic Setting
Review: An ODE is
\[
\left\{ \begin{array}{ll}
\frac{dy}{dt} = f(y, t) \\
y(0) = y_0
\end{array} \right.
\]
(1)
Eq. (1) can be viewed as the limit of finite difference equation:

\[ \Delta y = f(y, t) \Delta t \]

i.e.,

\[ y((k + 1)\delta t) - y(k\delta t) = f(y(k\delta t), k\delta t) \delta t \]

which is Taylor expansion, formally, we write this as

\[ dy = f(y, t) dt \]

Next, if we know \( dy = f(y, t) dt \), what ODE does \( z = A(y(t)) \) satisfy? This is answered by the chain rule:

\[ \frac{dz}{dt} = \frac{dA}{dy} \frac{dy}{dt} = \frac{dA}{dy} f(y, t) \]

i.e.,

\[ dz = A'f(y, t) dt \]

where

\[ A' \equiv \frac{dA}{dy} \]

Again, this can be viewed as the Taylor expansion:

\[ \Delta z = z(t + \Delta t) - z(t) = A(y(t) + \Delta y) - A(y(t)) \]
\[ = A'(y) \Delta y + O(\Delta y)^2 \]
\[ = A'(y) f(y, t) \Delta t + O(\Delta t)^2 \]

if

\[ (\Delta y)^2 \leq c(\Delta t)^2. \]

1.2.2 Stochastic Setting

Now we turn to the stochastic setting:

1. \( dy = \sigma dW \), by this short hand, we mean:

\[ y(t) = y(0) + \sigma \int_0^t dW \]
\[ = y(0) + \sigma W(t) \]

since \( W(0) = 0 \).
2. $dy = g(t) \, dW$, we mean
   
   $$y(t) = y(0) + \int_0^t g(s) \, dW(s)$$

   the question here is what is $\int_0^t g(s) \, dW(s)$? This is defined as

   $\int_0^t g(s) \, dW(s) = \lim_{\Delta \tau \to 0} \sum g(\tau_i) (W(\tau_{i+1}) - W(\tau_i))$

3. $dy = g(W, t) \, dW$ is interpreted as
   
   $$y(t) = y(0) + \int_0^t g(W(s), s) \, dW(s)$$

   where the integral $\int_0^t g(W(s), s) \, dW(s)$ is defined as the limit:

   $$\int_0^t g(W(s), s) \, dW(s) = \lim_{\Delta \tau \to 0} \sum g(W(\tau_i), \tau_i) (W(\tau_{i+1}) - W(\tau_i)).$$

   It is crucial that the dependence of $g$ on the discrete time is the beginning point of the interval $[\tau_i, \tau_{i+1}]$ and $W$ is incremented forward — Unlike the deterministic case, where the limit do not depend on the choice of discrete time between $g$ and $f$ for $\int gdf$, in the stochastic setting, the limit

   $$\lim_{\Delta \tau \to 0} \sum g(W(\tau_i'), \tau_i') (W(\tau_{i+1}) - W(\tau_i)),$$

   where

   $$\tau_i' \equiv \tau_i + \lambda(\tau_{i+1} - \tau_i),$$

   depends on $\lambda$. This leads to the concept of nonanticipatoriness, i.e., $g(W(\tau_i), \tau_i)$ does not contain the information of $dW = W(\tau_{i+1}) - W(\tau_i)$.

4. For any $g(W(s), s)$, we have
   
   $$\mathbb{E} \int_0^t g(W(s), s) \, dW(s) = 0$$

   This can be easily seen as follows:

   $\mathbb{E} \int_0^t g(W(s), s) \, dW(s) = \lim \sum \mathbb{E}[g(W(\tau_i), \tau_i) (W(\tau_{i+1}) - W(\tau_i))]$

   $= \lim \sum \mathbb{E}[g(W(\tau_i), \tau_i)] \mathbb{E}[(W(\tau_{i+1}) - W(\tau_i))]$

   (due to nonanticipatoriness, $g(W, s)$ and $dW$ are independent)

   $= \lim \sum \mathbb{E}[g(W(\tau_i), \tau_i)] \cdot 0$

   $= 0$ for any $g(W, s)$
5. The convergence we discussed above is the so-called mean-square convergence, i.e.,

$$\lim_{n \to \infty} E \left[ (X_n - X)^2 \right] = 0$$

6. Adding a drift, we have

$$dy = \mu (t) \, dt + \sigma (t) \, dW$$

which is merely a shorthand for

$$y (t) = y (0) + \int_0^t \mu (s) \, ds + \int_0^t \sigma (s) \, dW (s)$$

1.2.3 Ito’s Lemma

Now we are ready to answer the following question: If

$$dy = f \, dt + g \, dW$$

and

$$z = A (y)$$

what is the stochastic differential equation that \( z \) satisfies? Different from the answer we gave in the deterministic setting, the answer is given by the famous Ito’s Lemma:

$$dz = A' (y) \, dy + \frac{1}{2} A'' (y) \, g^2 \, dt$$

$$= \left( A' (y) \, f + \frac{1}{2} A'' (y) \, g^2 \right) \, dt + A' (y) \, g \, dW.$$

This can be heuristically justified as follows:

Taylor expansion of \( A (y) \) is

$$\Delta A = A' (y) \, \Delta y + \frac{1}{2} A'' (y) \, (\Delta y)^2 + \text{error of order } |\Delta y|^3$$

$$= A' (y) \, (g \Delta W + f \Delta t) + \frac{1}{2} A'' (y) \, g^2 \, (\Delta W)^2$$

$$+ \text{error of order } |\Delta y|^3, |\Delta W| \Delta t, |\Delta t|^2$$

Note that

$$E \left[ (\Delta W)^2 \right] = \Delta t$$

$$E \left[ |\Delta W| \right] \sim \Delta t^{1/2}$$

therefore,

$$|\Delta W| |\Delta t| \sim \Delta t^{3/2}$$

To order \( O (\Delta t) \), we have

$$dA = A' (y) \, (g dW + f dt) + \frac{1}{2} A'' (y) \, g^2 \, (dW)^2$$
Since, $E[(\Delta W)^2] = \Delta t$, and $(\Delta W)^2$ behaves more and more “deterministic” — like $\Delta t$, as $\Delta t \to 0$. In the mean-square sense, it turns out that we can replaced $(dW)^2$ by $dt$ in the expression, leading to

$$dA = \left(A'(y) f + \frac{1}{2} A''(y) g^2\right) dt + A'(y) g dW.$$  

A more general form of Ito’s lemma is

$$z = A(y,t),$$
$$dy = f dt + g dW,$$
$$dz = \frac{\partial A}{\partial y} dy + \frac{\partial A}{\partial t} dt + \frac{1}{2} \frac{\partial^2 A}{\partial y^2} g^2 dt$$

$$= \left(\frac{\partial A}{\partial t} + \frac{\partial A}{\partial y} f + \frac{1}{2} g^2 \frac{\partial^2 A}{\partial y^2}\right) dt + \frac{\partial A}{\partial y} g dW$$

Common rules for computation:

1. $E[(\Delta W)^2]$ is replaced by $dt$;
2. Order count:

$$(\Delta W)^n \sim \Delta t^{\frac{n}{2}}$$

We have the following nemonic form of Ito’s formula:

For

$$dy = f(t,y) dt + g(t,y) dW$$

and $z = A(y,t)$, we have

$$dz = \frac{\partial A}{\partial t} dt + \frac{\partial A}{\partial y} dy + \frac{1}{2} \frac{\partial^2 A}{\partial x^2} (dy)^2$$

with the replacement rule:

$$(dt)^2 = 0$$
$$dt dW = 0$$
$$(dW)^2 = dt$$
1.2.4 Examples:

1. Find the SDE for stock price process if \( S = S_0e^y \) and

\[
don y = \mu dt + \sigma dW
\]

Since

\[
y = y_0 + \mu t + \sigma W(t)
\]

\[
S_t = S_0e^{\mu t + \sigma W(t)}, \quad S_0 = e^{y_0}
\]

\[
dS = \frac{\partial S}{\partial y} dy + \frac{1}{2} \frac{\partial^2 S}{\partial y^2} \sigma^2 dt
\]

\[
= S_0e^y (\mu dt + \sigma dW) + \frac{S_0}{2} e^y \sigma^2 dt
\]

\[
= S (\mu dt + \sigma dW) + \frac{1}{2} \sigma^2 S dt
\]

\[
\therefore \quad \frac{dS}{S} = \left( \mu + \frac{1}{2} \sigma^2 \right) dt + \sigma dW
\]

which is called a geometric Brownian motion. Note that for a stock price satisfying the risk-neutral process, we have

\[
S_t = S_0e^{(r - \frac{1}{2} \sigma^2)t + \sigma W(t)}
\]

therefore,

\[
\frac{dS}{S} = \left( r - \frac{1}{2} \sigma^2 + \frac{1}{2} \sigma^2 \right) dt + \sigma dW
\]

\[
= r dt + \sigma dW
\]

2. Find the SDE for the forward price if

\[
dS = S \left( \mu + \frac{1}{2} \sigma^2 \right) dt + \sigma S dW
\]

Since

\[
F = Se^{r(T-t)}, \quad t < T
\]

\[
\therefore \quad dF = \frac{\partial F}{\partial S} dS + \frac{\partial F}{\partial t} dt + \frac{1}{2} \frac{\partial^2 F}{\partial S^2} (S\sigma)^2 dW
\]

\[
= e^{r(T-t)} dS - r Se^{r(T-t)} dt + 0 \cdot dt
\]

\[
= Se^{r(T-t)} \frac{dS}{S} - r (e^{r(T-t)}) dt
\]

\[
= F \left[ \left( \mu + \frac{1}{2} \sigma^2 \right) dt + \sigma dW \right] - r F dt
\]
Hence,
\[ dF = \left( \mu - r + \frac{1}{2} \sigma^2 \right) F dt + \sigma F dW \]
which is also a geometric Brownian motion.

Note that different books have different notational conventions, e.g., Hull’s book uses \( dS = \mu S dt + \sigma S dW \). Don’t copy formulas with checking the convention.

3. Compute \( \mathbb{E} [e^{\alpha W(t)}] \)
\[ Z(t) \equiv e^{\alpha W(t)} \]
\[ dZ(t) = \frac{1}{2} \alpha^2 e^{\alpha W(t)} dt + \alpha e^{\alpha W(t)} dW \]
therefore, \( Z(t) \) satisfies the SDE:
\[ dZ(t) = \frac{1}{2} \alpha^2 Z(t) dt + \alpha Z(t) dW(t) \quad \text{with } Z(0) = 1 \]

or
\[ Z(t) = 1 + \frac{1}{2} \alpha^2 \int_0^t Z(s) ds + \alpha \int_0^t Z(s) dW(s) \]
\[ m(t) \equiv \mathbb{E} [Z(t)] \]
therefore,
\[ m = 1 + \frac{1}{2} \alpha^2 \int_0^t m(s) ds \]
i.e.,
\[ \frac{dm}{dt} = \frac{1}{2} \alpha^2 m, \quad \text{with } m(0) = 1 \]
\[ \implies \mathbb{E} [e^{\alpha W(t)}] = \mathbb{E}[Z(t)] = m(t) = e^{\frac{1}{2} \alpha^2 t} \]

1.3 The Black-Scholes PDE

The Black-Scholes PDE is
\[ \frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} - rV = 0, \]
where
\[ V = V(S(t), t) \] — the value of a contingent claim at time \( t \),
describes the time dynamics of the value of a contingent claim. The BS PDE is solved with the following final condition:
\[ V(S(T), T) = f(S_T) \]
where \( f(S_T) \) is the payoff of the contingent claim at maturiy \( T \).
Since we know a European option with payoff \( f(S_T) \) has the value at time \( t \):

\[
V(S(t), t) = e^{-r(T-t)} \int_{-\infty}^{\infty} f(S_0 e^x) \frac{1}{\sqrt{2\pi\sigma^2(T-t)}} e^{-\frac{[x-(r-\frac{1}{2}\sigma^2)r/2]}{2\sigma^2(T-t)}} dx
\]

it must satisfy the BS PDE.

Note that

1. The BS integral formula is the continuous-time limit analogue of summing over all paths

2. Solving the BS PDE is the continuous-time analogue of working backward through the tree by starting from the final data

### 1.3.1 Derivation

Now we derive the BS PDE using Ito’s lemma. The central idea involved is a hedging strategy — which, of course, implies that our contingent claim is replicatable, as a consequence of the completeness of the Black-Scholes market (a deep result beyond our course).

Consider the hedging portfolio:

at time \( t \), \( \Pi = V - \Delta S \),

where \( \Delta \) is the number of shares of stock in this portfolio. At time \( t + dt \), the increment of the value of the portfolio is

\[
d\Pi = \Pi(t + dt) - \Pi(t) = [V(t + dt) - \Delta S(t + dt)] - [V(t) - \Delta S(t)]
\]

the number of shares are fixed from \( t \) to \( t + dt \)

\[
= dV - \Delta dS = \left[ \frac{\partial V}{\partial S} dS + \frac{\partial V}{\partial t} dt + \frac{1}{2} \frac{\partial^2 V}{\partial S^2} \sigma^2 S^2 dt \right] - \Delta dS
\]

\[
= \left( \frac{\partial V}{\partial S} - \Delta \right) \sigma S dW + \left( \frac{\partial V}{\partial S} - \Delta \right) \left( \mu + \frac{1}{2} \sigma^2 \right) dt
\]

\[
+ \left( \frac{\partial V}{\partial t} + \frac{1}{2} \frac{\partial^2 V}{\partial S^2} \sigma^2 S^2 \right) dt
\]

Note that the number of shares of the stock in this portfolio is not changed from time \( t \) to time \( t + dt \) since no trading is involved. To remove uncertainty, i.e, eliminating \( dW \), we choose

\[
\Delta = \frac{\partial V}{\partial S}
\]
This leads to

\[ d\Pi = dV - \Delta dS = \left( \frac{\partial V}{\partial t} + \frac{1}{2} \frac{\partial^2 V}{\partial S^2} \sigma^2 S^2 \right) dt \]

i.e., the changing of the value of the portfolio is deterministic — there is no risk involved in this portfolio, then the no-arbitrage principle demands that a deterministic return must be at the risk-free rate, therefore,

\[ d\Pi = r\Pi dt = r(V - \Delta S) dt \quad \text{with} \quad \Delta = \frac{\partial V}{\partial S} \]

which yields

\[ \left( \frac{\partial V}{\partial t} + \frac{1}{2} \frac{\partial^2 V}{\partial S^2} \sigma^2 S^2 \right) dt = r \left( V - \frac{\partial V}{\partial S} S \right) dt \]

i.e.,

\[ \frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} - rV = 0, \]

which is the Black-Scholes PDE.

Note that \( \Delta = \frac{\partial V}{\partial S} \), the number of shares in the hedging portfolio, is consistent with what we already know in the \( \Delta \)-hedging strategy, i.e., the number of share is the ratio of the change of the option value to the change of the stock price.

### 1.3.2 The solution of BS PDE and the risk-neutral pricing formula

Now we prove that the solution of the BS PDE gives our risk-neutral pricing formula:

\[ V = V(S_t, t) = e^{-r(T-t)} \mathbb{E}[f(S_T)] \]

with the risk-neutral process:

\[ S_T = S_t e^{(r - \frac{1}{2}\sigma^2)(T-t) + \sigma(W(T) - W(t))} \]

Note that this process satisfies

\[ dS_t = rS_t dt + \sigma S_t dW_t, \quad t \leq T \]

Define

\[ U(S, t) = e^{r(T-t)} V(S, t) \]
i.e., to remove the discount factor from the expression of \( V(S,t) \). Applying Ito’s formula:

\[
\begin{align*}
\frac{dU}{dS} &= \frac{\partial U}{\partial S} dS + \frac{\partial U}{\partial t} dt + \frac{1}{2} \frac{\partial^2 U}{\partial S^2} \sigma^2 S^2 dt \\
&= e^{r(T-t)} \frac{\partial V}{\partial S} dS + (-r) e^{r(T-t)} V dt \\
&\quad + e^{r(T-t)} \frac{\partial V}{\partial t} dt + \frac{1}{2} e^{r(T-t)} \frac{\partial^2 V}{\partial S^2} \sigma^2 S^2 dt
\end{align*}
\]

(\because ds = rSdt + \sigma SdW for risk-neutral process)

\[
\begin{align*}
&= e^{r(T-t)} \left\{ -rV + \frac{\partial V}{\partial t} + rS \frac{\partial V}{\partial S} + \frac{1}{2} \frac{\partial^2 V}{\partial S^2} \sigma^2 S^2 \right\} dt + \sigma S \frac{\partial V}{\partial S} dW \\
&\quad \left\{ \begin{array}{l}
\quad = 0 \quad (\because \text{BS PDE=0, i.e., } V \text{ is a solution of it})
\end{array} \right.
\end{align*}
\]

therefore,

\[
\begin{align*}
\frac{dU}{dS} &= e^{r(T-t)} \sigma S \frac{\partial V}{\partial S} dW \\
&= \sigma S \frac{\partial U}{\partial S} dW
\end{align*}
\]

i.e., the normalized \( V, U \), is a martingale. Hence,

\[
U(t') - U(t) = \int_t^{t'} e^{r(t'-\tau)} \sigma S(\tau) \frac{\partial V}{\partial S}(S(\tau), \tau) dW(\tau), \quad t' > t
\]

an Ito integral, whose expectation vanishes

therefore,

\[
\mathbb{E}(U(t')) = \mathbb{E}(U(t)) \quad (2)
\]

For \( t' = T \), we have then

\[
\mathbb{E}(U(t')) = \mathbb{E}(U(T))
\]

(\text{using the final data for BS pde} : \quad U(T) = e^{r(T-T)} V(T) = f(S_T))

\[
= \mathbb{E}(f(S_T))
\]

Moreover, since \( U(t) \) is deterministically known with certainty at time \( t \), we have

\[
\mathbb{E}(U(t)) = U(t) = e^{r(T-t)} V(S(t), t)
\]

therefore, from Eq. (2), we obtain

\[
e^{r(T-t)} V(S(t), t) = \mathbb{E}(f(S_T))
\]

i.e.,

\[
V = e^{-r(T-t)} \mathbb{E}(f(S_T))
\]

Note that
1. the expectation is taken with respect to a Brownian motion $dW$, under this brownian motion, the stock price is $dS = rSdt + \sigma SdW$ — the risk-neutral process — is used in derivation, thus

$$\mathbb{E}(\cdot) = \mathbb{E}_{RN}(\cdot)$$

2. Tradeable derivatives

(a) Any solution of the BS PDE is a theoretical price of a derivative that can be traded without arbitrage.

(b) If a price $V(S_t, t)$ does not satisfy the BS PDE, then, it cannot be a price of some derivative in the absence of arbitrage.

  e.g.,
  i. $V = e^S$, this $V$ does not satisfy the BS pde (verify this!), then $V = e^S$ cannot be the price of some derivative without arbitrage. In other words, it cannot be replicated with a self-financing portfolio.

  ii. $\frac{e^{(\sigma^2-2r)(T-t)}}{S}$ satisfies the BS pde (verify this!), then it can be a price of some derivative (i.e, the payoff is $\frac{1}{S_T}$ at maturity $t = T$.)