

1 Itô's Calculus

Throughout, we use B_t to denote the standard Brownian motion (BM). Let

$$dB_s := B_{s+ds} - B_s$$

denote an increment of the BM (with $ds > 0$). We also use $N(\mu, \sigma^2)$ to denote a normal distribution with mean μ and variance σ^2 .

Recall some of the key properties of BM: (i) $B_0 = 0$; (ii) independent increments, i.e., dB_s and dB_t are independent, for any $s + ds \leq t$; (iii) stationary increments, i.e., dB_s follows a normal distribution $N(0, ds)$. Note this last distribution depends only on the length of the increment, not on when it starts (hence, “shift invariant”, or stationary). Also note that the *variance* is proportional to the length of the increment. The BM is a Markov process, with a continuous trajectory over time.

We state without formal proof the following result:

$$[dB_s]^2 = ds, \quad \text{as } ds \rightarrow 0. \quad (1)$$

Note that whereas $E([dB_s]^2) = ds$, we have, with Z denoting the standard normal variate,

$$\text{Var}([dB_s]^2) = \text{Var}(ds \cdot Z^2) = (ds)^2 [E(Z^4) - E^2(Z^2)] = 2(ds)^2.$$

(Note $E(Z^4) = 3$.) Therefore, the relation in (1) appeals to intuition: the random variable on the left hand side has a variance that is a higher-order infinitesimal than its mean. Hence, as $ds \rightarrow 0$, the random variable becomes deterministic.

In general, the multiplication rules in Table 1, the so-called “box algebra”¹ is useful when it comes to taking derivatives on functions that involve Brownian motion. (The zero’s in the table should be read as higher-order infinitesimals w.r.t. dt .) And this is essentially what leads to *Itô’s calculus*.

To motivate, consider the ordinary calculus. Suppose x_t is a deterministic function of time, or, a deterministic “path/trajectory”. Write $dx_t = \dot{x}dt$, where $\dot{x}$ denotes the derivative of x over time. Consider a smooth function f . We have

$$df(x_t) = f'(x_t)dx_t.$$

¹J.M. Steele, *Stochastic Calculus and Financial Applications*, Springer-Verlag, New York, 2001.

$\cdot$	dt	dB_t
dt	0	0
dB_t	0	dt

Table 1: “Box Algebra”

More precisely,

$$df(x_t) = f'(x_t)dx_t + \frac{1}{2}f''(x_t)(dx_t)^2 + \cdots,$$

with the higher-order terms vanishing when $dt \rightarrow 0$. For instance, $(dx_t)^2 = \dot{x}^2(dt)^2$.

Now, if x_t is replaced by the BM, B_t , then in view of (1), we must include the second-derivative term, since $(dB_t)^2 = dt$, i.e., at the same order as dt . This results in what is known as *Itô's formula*:

$$df(B_t) = f'(B_t)dB_t + \frac{1}{2}f''(B_t)dt. \quad (2)$$

Taking integral on both sides, we have another form of Itô's formula:

$$f(B_t) = f(0) + \int_0^t f'(B_s)dB_s + \frac{1}{2} \int_0^t f''(B_s)ds. \quad (3)$$

As an example, consider $f(X_t)$, where

$$X_t = \mu t + \sigma B_t. \quad (4)$$

(X_t is the generalized Wiener process, or Brownian motion with drift.) Following (2), we have

$$df(X_t) = f'(X_t)dX_t + \frac{1}{2}f''(X_t)(dX_t)^2 + o((dX_t)^2). \quad (5)$$

Since

$$\begin{aligned} (dX_t)^2 &= (\mu dt + \sigma dB_t)^2 \\ &= \mu^2(dt)^2 + 2\mu\sigma(dt)(dB_t) + \sigma^2(dB_t)^2 \\ &= \sigma^2 dt, \end{aligned} \quad (6)$$

where on the last line we have ignored terms that are of higher order w.r.t. dt . Substituting (6) into (5) and omitting higher-order terms, we have

$$\begin{aligned} df(X_t) &= f'(X_t)dX_t + \frac{1}{2}\sigma^2 f''(X_t)dt \\ &= [\mu f'(X_t) + \frac{1}{2}\sigma^2 f''(X_t)]dt + \sigma f'(X_t)dB_t. \end{aligned} \quad (7)$$

More generally, consider a bivariate function $f(t, x)$. The formuals in (2) and (3) extend to:

$$df(t, B_t) = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dB_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} dt, \quad (8)$$

and

$$f(t, B_t) = f(0, 0) + \int_0^t f'_t(s, B_s) ds + \int_0^t f'_x(s, B_s) dB_s + \frac{1}{2} \int_0^t f''_{xx}(s, B_s) ds, \quad (9)$$

where $f'_t := \frac{\partial f}{\partial t}$, $f'_x := \frac{\partial f}{\partial x}$, and $f''_{xx} := \frac{\partial^2 f}{\partial x^2}$.

Consider, for example, $Y_t = \exp(\mu t + \sigma B_t)$. Here, $f(t, x) = \exp(\mu t + \sigma x)$. Since

$$f'_t = \mu \cdot f, \quad f'_x = \sigma \cdot f, \quad f''_{xx} = \sigma^2 \cdot f,$$

we have

$$dY_t = Y_t \cdot [(\mu + \frac{1}{2}\sigma^2)dt + \sigma dB_t].$$

Alternatively, we can also write $Y_t = e^{X_t}$, with X_t following (4), and make use of (7) to derive the same expression as above. (In this case, $f(x) = e^x = f' = f''$.)

2 Stock-Price Dynamics: Geometric Brownian Motion

Let S_t denote the price of a certain stock at time t . We model the dynamics of S_t as *geometric Brownian motion* (GBM):

$$\frac{dS_t}{S_t} = \mu dt + \sigma dB_t. \quad (10)$$

[It would be useful to relate the above to Model-I discussed earlier in class:

$$\frac{S_{t+\Delta} - S_t}{S_t} = \mu\Delta + \sigma\sqrt{\Delta}Z,$$

where Δ is a small time increment, and Z is the standard normal variate.]

Now, consider $\ln S_t$. We want to derive $d(\ln S_t)$. Note the function in question is $f(x) = \ln x$, and $f'(x) = \frac{1}{x}$ and $f''(x) = -\frac{1}{x^2}$. We have, following Itô's calculus [cf. (5)],

$$\begin{aligned} d(\ln S_t) &= \frac{dS_t}{S_t} - \frac{1}{2} \frac{(dS_t)^2}{S_t^2} \\ &= \mu dt + \sigma dB_t - \frac{1}{2} (\mu dt + \sigma dB_t)^2 \\ &= \mu dt + \sigma dB_t - \frac{1}{2} \sigma^2 dt \\ &= \nu dt + \sigma dB_t, \end{aligned}$$

where the second equality follows from (10), the third follows from (6), and in the last one we denote $\nu := \mu - \frac{1}{2}\sigma^2$ as before. Taking integral, from 0 to T , on both sides above, we have

$$\ln S_T - \ln S_0 = \nu T + \sigma B_T,$$

i.e., $S_T = S_0 e^{\nu T + \sigma B_T}$; or, switch to a more generic t :

$$S_t = S_0 e^{\nu t + \sigma B_t}. \quad (11)$$

In particular, we know from above that for any given time T , S_T follows a lognormal distribution:

$$\ln S_T \sim N(\ln S_0 + (\mu - \sigma^2/2)T, \sigma^2 T). \quad (12)$$

[Recall from Model-II of an earlier lecture,

$$\frac{S_{t+\Delta}}{S_t} = e^{\nu\Delta + \sigma\sqrt{\Delta}Z},$$

we have reached the same conclusion of a lognormal distribution for S_T .]

It is important to keep in mind that whereas the GBM in (10) characterizes the *dynamics* of the stock price over time, the lognormal distribution above only specifies the distribution of the stock price at a single time point T . That is, the former carries much more information than the latter.

As another example of applying Itô's calculus, let us take derivative on S_t from (11), replacing T by t . The function involved is $f(x) = e^x$, and $f'(x) = f''(x) = f(x)$. Hence, we have

$$\begin{aligned} dS_t &= S_t[(\nu dt + \sigma dB_t) + \frac{1}{2}(\nu dt + \sigma dB_t)^2] \\ &= S_t[(\nu dt + \sigma dB_t) + \frac{1}{2}\sigma^2 dt] \\ &= S_t(\mu dt + \sigma dB_t), \end{aligned}$$

recovering (10).

It will be useful to remember from now on the two equivalent forms of the GBM stock-price model in (10) and in (11). In particular, notice the two different drift parameters in the two forms, μ in the differential form and ν in the exponential form.

3 Option Pricing Formulae: Risk-Neutral Valuation

Consider a European call option with strike price K and maturity T . Let S_t denote the price of the underlying stock. Let r denote the risk-neutral interest rate. The value of the option, at

time $t = 0$, is:

$$c = e^{-rT} \hat{\mathbb{E}}[S_T - K]^+, \quad (13)$$

where $[x]^+ := \max\{x, 0\}$. That is, the value of the option is nothing but the present value of its expected payoff.

The key here is: the expectation, $\hat{\mathbb{E}}$, is w.r.t. the following distribution of the S_T :

$$\ln S_T \sim N(\ln S_0 + (r - \sigma^2/2)T, \sigma^2 T) := N(m, \eta^2). \quad (14)$$

That is, the intrinsic rate of the stock, μ in (10) and (12), is replaced by the risk-neutral interest rate r ; and in this sense, the valuation in (13) is termed *risk-neutral*. Recall, this is motivated by what we learned from analyzing the binomial tree model.

What remains is to derive the expectation in (13). To simplify notation, below we shall drop the “hat” from $\hat{\mathbb{E}}$, with the understanding that the distribution in question is the one in (14). Our starting point is to note that the expectation in (13) can be rewritten as follows:

$$\begin{aligned} \mathbb{E}[S_T - K]^+ &= \mathbb{E}[(S_T - K)\mathbf{1}(S_T \geq K)] \\ &= \mathbb{E}[S_T \mathbf{1}(S_T \geq K)] - K\mathbb{P}(S_T \geq K), \end{aligned} \quad (15)$$

where $\mathbf{1}$ denotes the indicator function, and we have used the fact that $\mathbb{E}\mathbf{1}(A) = \mathbb{P}(A)$ for any event A .

From (14), with m and η denoting the mean and variance of the normal distribution in question, we can write

$$S_T = e^{m+\eta Z}. \quad (16)$$

Hence,

$$\begin{aligned} \mathbb{P}(S_T \geq K) &= \mathbb{P}(m + \eta Z \geq \ln K) \\ &= \mathbb{P}(Z \geq \frac{\ln K - m}{\eta}) \\ &= \bar{\Phi}(\frac{\ln K - m}{\eta}) \\ &= \Phi(\frac{m - \ln K}{\eta}) := \Phi(d_2). \end{aligned} \quad (17)$$

[Recall, $\Phi(x)$ denotes the distribution function of Z , and $\bar{\Phi}(x) = 1 - \Phi(x) = \Phi(-x)$.]

Next, using the expression in (16) once again, we have

$$\begin{aligned}
& \mathbb{E}[S_T \mathbf{1}(S_T \geq K)] \\
&= \frac{1}{\sqrt{2\pi}} \int_{m+\eta x \geq \ln K} e^{m+\eta x} \cdot e^{-x^2/2} dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{x \geq -d_2} e^{-(x-\eta)^2/2} \cdot e^{m+\eta^2/2} dx \\
&= e^{m+\eta^2/2} \mathbb{P}(Z + \eta \geq -d_2).
\end{aligned}$$

Note that in the second equality we performed a standard completion of square, and d_2 follows its definition in (17). The third equality follows from recognizing the density function of $Z + \eta$ involved in the second integral.

Hence, if we write $d_1 := d_2 + \eta$, and notice that

$$e^{m+\eta^2/2} = S_0 e^{rT}$$

follows from (14), then we have

$$\mathbb{E}[S_T \mathbf{1}(S_T \geq K)] = S_0 e^{rT} \mathbb{P}(Z \geq -d_1) = S_0 e^{rT} \Phi(d_1). \quad (18)$$

Therefore, combining (17) and (18) into (15), and further into (13), we have

$$c = S_0 \Phi(d_1) - K e^{-rT} \Phi(d_2), \quad (19)$$

where d_1 and d_2 , following their earlier definitions, can be written explicitly as follows:

$$d_1 := d_2 + \sigma\sqrt{T}, \quad d_2 := \frac{\ln(S_0/K) + (r - \sigma^2/2)T}{\sigma\sqrt{T}}. \quad (20)$$

Next, consider the put option. With risk-neutral valuation, it can be priced as [cf. (13)]

$$p = e^{-rT} \hat{\mathbb{E}}[K - S_T]^+. \quad (21)$$

Applying the following simple identity, which holds for any a and b values,

$$(a - b)^+ - (b - a)^+ = a - b, \quad (22)$$

to (13) and (21), we have

$$c - p = e^{-rT} \hat{\mathbb{E}}[S_T - K] = e^{-rT} [\hat{\mathbb{E}}(S_T) - K].$$

Since

$$\hat{\mathbb{E}}(S_T) = \mathbb{E}[S_0 e^{(r-\sigma^2/2)T + \sigma\sqrt{T}Z}] = S_0 e^{(r-\sigma^2/2)T} \mathbb{E}[e^{\sigma\sqrt{T}Z}] = S_0 e^{rT},$$

we have the following *put-call parity* relation,

$$c - p = S_0 - K e^{-rT}. \quad (23)$$

Substituting (19) into (23) leads to:

$$p = -S_0 \Phi(-d_1) + K e^{-rT} \Phi(-d_2), \quad (24)$$

with d_1 and d_2 still following (20).

4 The PDE

Let $f(t, S_t)$ denote the value (price) of the option at time t , when the underlying stock price is S_t . Note that as yet, we do not make a distinction upon whether the option is a call or a put. Also note that here the *present* time is set at a more general t , rather than 0. Assume S_t follows the geometric Brownian motion model in (10). Let $\frac{\partial f}{\partial S_t}$ and $\frac{\partial f}{\partial t}$ denote the partial derivatives of f w.r.t. the two variables, and let $\frac{\partial^2 f}{\partial S_t^2}$ denote the second-order partial derivative of f w.r.t. to S_t .

To value the option – more specifically, to derive a *partial differential equation* (pde) that f must satisfy – we follow two steps that are analagous to the (discrete-time) binomial-tree model. First, we construst a *replicating portfolio*, which consists of $\frac{\partial f}{\partial S_t}$ shares of the stock and -1 share of the option. Let Π_t denote the value of the portfolio at t . We have

$$\Pi_t = -f(t, S_t) + \frac{\partial f}{\partial S_t} \cdot S_t. \quad (25)$$

The change of the portfolio value over $(t, t + dt)$ is:

$$d\Pi_t = -df(t, S_t) + \frac{\partial f}{\partial S_t} \cdot dS_t. \quad (26)$$

For the first term on the right hand side above, applying Itô's calculus and making use of the stock dynamics in (10), we have

$$\begin{aligned} df(t, S_t) &= \frac{\partial f}{\partial S_t} \cdot dS_t + \frac{1}{2} \frac{\partial^2 f}{\partial S_t^2} (dS_t)^2 + \frac{\partial f}{\partial t} dt \\ &= \frac{\partial f}{\partial S_t} \cdot dS_t + \frac{1}{2} \frac{\partial^2 f}{\partial S_t^2} (S_t^2 \sigma^2 dt) + \frac{\partial f}{\partial t} dt. \end{aligned}$$

Substituting the above into (26), we have

$$d\Pi_t = -\left(\frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 f}{\partial S_t^2} + \frac{\partial f}{\partial t}\right) dt. \quad (27)$$

Note from the above derivation that the choice of $\frac{\partial f}{\partial S_t}$ shares of stock in the replicating portfolio cancels out the term, $\frac{\partial f}{\partial S_t} dS_t$, which involves randomness. What remains, the right hand side of (27), are non-random.

Second, we apply the *no-arbitrage* argument: As $d\Pi_t$ involves no randomness, i.e., no risk, the portfolio's rate of return over $(t, t + dt)$ should be no more and no less than the risk-free interest rate, r . That is,

$$d\Pi_t = \Pi_t r dt. \quad (28)$$

Substituting (25) and (27) into the above equation, cancelling out dt on both sides and rearranging terms, we have

$$\frac{\partial f}{\partial t} + rS_t \frac{\partial f}{\partial S_t} + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 f}{\partial S_t^2} = rf(t, S_t). \quad (29)$$

This is the Black-Scholes pde that determines the price (f) of the option at time t

At maturity $t = T$, depending on whether the option is a call or a put, we have

$$f(T, S_T) = [S_T - K]^+, \quad \text{or} \quad f(T, S_T) = [K - S_T]^+, \quad (30)$$

which are nothing but the payoff functions (at maturity). These are called the *boundary conditions* to the pde. They determine whether f is the price of a call or a put.

5 Solutions to the PDE, the Greek Letters

We now verify that the formula in §3 derived via risk-neutral valuation satisfy the pde and its boundary condition. We focus on the call option, the put being completely analogous.

To start with, we modify the formulas in (19) and (20) to allow the present time to be t instead of 0, which amounts to replacing 0 by t , and T by $T - t$:

$$c(t, S_t) = S_t \Phi(d_1) - K e^{-r(T-t)} \Phi(d_2), \quad (31)$$

where

$$d_1 := d_2 + \sigma \sqrt{T-t}, \quad d_2 := \frac{\ln(S_t/K) + (r - \sigma^2/2)(T-t)}{\sigma \sqrt{T-t}}. \quad (32)$$

First, we verify the boundary condition, the first equation in (30). Letting $t \rightarrow T$ in (32), we have three cases:

- (i) $d_1 = d_2 \rightarrow +\infty$, if $S_T > K$; hence, $\Phi(d_1) = \Phi(d_2) = 1$, and $c(T, S_T) = S_T - K$;

(ii) $d_1 = d_2 \rightarrow -\infty$, if $S_T < K$; hence, $\Phi(d_1) = \Phi(d_2) = 0$ and $c(T, S_T) = 0$;

(iii) if $S_T = K$, then since $d_1 = d_2$ when $t \rightarrow T$, we have $c(T, S_T) = 0$.

Hence, we have $c(T, S_T) = [S_T - K]^+$; i.e., the boundary condition in (30) is indeed satisfied.

To verify that the c function satisfies the pde in (29), we need to derive the three derivatives involved in the pde. The following relation will prove to be useful:

$$S_t \phi(d_1) = K e^{-(T-t)} \phi(d_2). \quad (33)$$

To verify the above, notice that $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$, and hence,

$$\frac{\phi(d_1)}{\phi(d_2)} = \exp\left[-\frac{1}{2}(d_1^2 - d_2^2)\right];$$

furthermore, from $d_1 = d_2 + \sigma\sqrt{T-t} := d_2 + \eta$, we have

$$\frac{1}{2}(d_1^2 - d_2^2) = \eta d_2 + \eta^2/2 = \ln(S_t/K) + r(T-t),$$

where the second equality follows from d_2 in (32). Combining the last two displays yields

$$\frac{\phi(d_1)}{\phi(d_2)} = \exp[-\ln(S_t/K) - r(T-t)] = (K/S_t) e^{-r(T-t)},$$

which is the identity in (33).

Now we are ready to take (partial) derivatives on $c(t, S_t)$ from (31). We have,

$$\begin{aligned} \frac{\partial c}{\partial S_t} &= \Phi(d_1) + S_t \frac{\partial}{\partial S_t} \Phi(d_1) - K e^{-r(T-t)} \frac{\partial}{\partial S_t} \Phi(d_2) \\ &= \Phi(d_1) + S_t \phi(d_1) \frac{\partial d_1}{\partial S_t} - K e^{-r(T-t)} \phi(d_2) \frac{\partial d_2}{\partial S_t} \\ &= \Phi(d_1) := \Delta, \end{aligned} \quad (34)$$

where the last equality follows from (33) and the fact that

$$\frac{\partial}{\partial S_t}(d_1 - d_2) = \frac{\partial}{\partial S_t}(\sigma\sqrt{T-t}) = 0.$$

Taking partial derivative w.r.t. S_t on both sides of (34) yields:

$$\frac{\partial^2 c}{\partial S_t^2} = \phi(d_1) \frac{\partial d_1}{\partial S_t} = \phi(d_1) \frac{\partial d_2}{\partial S_t} = \frac{\phi(d_1)}{S_t \sigma \sqrt{T-t}} := \Gamma. \quad (35)$$

Next, we take partial derivative w.r.t. t on (31):

$$\frac{\partial c}{\partial t} = S_t \phi(d_1) \frac{\partial d_1}{\partial t} - r K e^{-r(T-t)} \Phi(d_2) - K e^{-r(T-t)} \phi(d_2) \frac{\partial d_2}{\partial t}.$$

Note the first and the third terms above can be combined, taking into account (33). This, along with the following:

$$\frac{\partial}{\partial t}(d_1 - d_2) = \frac{\partial}{\partial t}(\sigma\sqrt{T-t}) = \frac{-\sigma}{2\sqrt{T-t}},$$

yields

$$\frac{\partial c}{\partial t} = -\frac{\sigma S_t \phi(d_1)}{2\sqrt{T-t}} - rKe^{-r(T-t)}\Phi(d_2) := \Theta. \quad (36)$$

Finally, combining (34), (35) and (36), into the left hand side of (29), we have

$$\begin{aligned} & \Theta + rS_t\Delta + \frac{1}{2}\sigma^2 S_t^2 \Gamma \\ &= -\frac{\sigma S_t \phi(d_1)}{2\sqrt{T-t}} - rKe^{-r(T-t)}\Phi(d_2) + rS_t\Phi(d_1) + \frac{\sigma S_t \phi(d_1)}{2\sqrt{T-t}} \\ &= -rKe^{-r(T-t)}\Phi(d_2) + rS_t\Phi(d_1) \\ &= rc(t, S_t). \end{aligned} \quad (37)$$

Therefore, we have verified that the pricing formula in (31) does satisfy the pde in (29).

The partial derivatives, Δ and Θ , measure the sensitivity of the option w.r.t. the stock price and the time, respectively. (Note that Δ is positive and Θ negative.) More sensitivity measures are derived below:

$$\begin{aligned} \frac{\partial c}{\partial \sigma} &= S_t \phi(d_1) \frac{\partial d_1}{\partial \sigma} - Ke^{-r(T-t)} \phi(d_2) \frac{\partial d_2}{\partial \sigma} \\ &= S_t \phi(d_1) \frac{\partial}{\partial \sigma}(d_1 - d_2) \\ &= S_t \phi(d_1) \sqrt{T-t} := \mathcal{V}, \end{aligned} \quad (38)$$

where the second equality follows from (33); and, similar to the derivation of $\frac{\partial c}{\partial t}$ above:

$$\begin{aligned} \frac{\partial c}{\partial r} &= S_t \phi(d_1) \frac{\partial d_1}{\partial r} + K(T-t)e^{-r(T-t)}\Phi(d_2) - Ke^{-r(T-t)}\phi(d_2) \frac{\partial d_2}{\partial r} \\ &= K(T-t)e^{-r(T-t)}\Phi(d_2) := \rho. \end{aligned} \quad (39)$$

To summarize, we know the call option c is increasing and convex in the stock price S_t ($\Delta \geq 0$, $\Gamma \geq 0$), decreasing in time t ($\Theta \leq 0$), increasing in the volatility σ ($\mathcal{V} \geq 0$), and increasing in the interest rate r ($\rho \geq 0$).

Exercise. Derive all the Greek letters for the put option, making use of the put-call parity in (23).

6 Delta Hedging

[This section and the next one draw materials from Chapter 4 of *Stochastic Calculus for Finance II, Continuous-Time Models*, by S.E. Shreve, Springer, New York, 2004.]

Suppose we have an amount of capital, initially valued at V_0 . At any time t , we invest part of the capital in Δ_t shares of a stock (each share is valued at S_t); and the remaining part, $V_t - \Delta_t S_t$, in a money-market account (which returns at the risk-free rate r). We shall refer to this as the hedging portfolio.

The change of V_t , over the time interval $(t, t + dt)$, consists of two parts: (a) the capital gain (or loss) in stock, $\Delta_t dS_t$; and (b) the interest income from the money-market account, $r(V_t - \Delta_t S_t)dt$. Hence, we have

$$dV_t = \Delta_t dS_t + r(V_t - \Delta_t S_t)dt. \quad (40)$$

Now, suppose our initial capital is equal to the price of an option, $V_0 = f(0, S_0)$. If we can make the changes of our capital position follow the changes in the option price all the time, i.e.,

$$d(e^{-rt}V_t) = d(e^{-rt}f(t, S_t)), \quad \forall t \in (0, T); \quad (41)$$

then we will have $V_t = f(t, S_t)$ for all $t \in (0, T)$; and letting $t \rightarrow T$, we will also have $V_T = f(T, S_T)$, with the latter equal to either $(S_T - K)^+$ or $(K - S_T)^+$ depending on whether the option is a call or a put.²

Let us first derive the LHS of (41), applying Itô's formula. The functional in question is $g(t, x) = e^{-rt}x$; and $g'_t = -rg$, $g'_x = e^{-rt}$, and $g''_{xx} = 0$. Hence,

$$\begin{aligned} d(e^{-rt}V_t) &= e^{-rt}[-rV_t dt + dV_t] \\ &= e^{-rt}[-rV_t dt + \Delta_t dS_t + r(V_t - \Delta_t S_t)dt] \\ &= e^{-rt}[\Delta_t dS_t - r\Delta_t S_t dt]. \end{aligned} \quad (42)$$

Similarly,

$$\begin{aligned} d(e^{-rt}f(t, S_t)) &= e^{-rt}[-rf(t, S_t)dt + df(t, S_t)] \\ &= e^{-rt}[-rf dt + \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial S_t} \cdot dS_t + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 f}{\partial S_t^2} dt], \end{aligned} \quad (43)$$

where the expression for $df(t, S_t)$ follows from the derivation in §4, the equation above (27).

²What's described here is exactly the delta hedging mechanism described in Tables 13.1 and 13.2 in Hull (also refer to the spreadsheet). Note, in particular, that the total cumulative cost of the hedging is close to the option price at time zero. In other words, if the bank uses an initial capital that is equal to the option price to do the hedging, it will break even at maturity.

Hence, in order to have (41), we need to equate the coefficients of dS_t and dt in (42) and (43). That is,

$$\Delta_t = \frac{\partial f}{\partial S_t}, \quad (44)$$

$$-r\Delta_t S_t = -rf + \frac{\partial f}{\partial t} + \frac{1}{2}S_t^2\sigma^2\frac{\partial^2 f}{\partial S_t^2}; \quad (45)$$

the former being the definition of the Greek letter delta; the latter, the PDE in (29).

7 The PDE Revisited

In the last section, we have, through delta hedging, effectively re-derived the PDE that governs the option price $f(t, S_t)$. Here, we take another look at the PDE, and also bring out a few other important concepts and tools.

7.1 Two-Dimensional Itô's Formula

Consider two processes:

$$\begin{aligned} X_t &= \mu_1 t + \sigma_{11}B_t^1 + \sigma_{12}B_t^2, \\ Y_t &= \mu_2 t + \sigma_{21}B_t^1 + \sigma_{22}B_t^2; \end{aligned}$$

where B_t^1 and B_t^2 are two independent Brownian motions. Then,

$$\begin{aligned} dX_t &= \mu_1 dt + \sigma_{11}dB_t^1 + \sigma_{12}dB_t^2, \\ dY_t &= \mu_2 dt + \sigma_{21}dB_t^1 + \sigma_{22}dB_t^2. \end{aligned}$$

Here, the “box algebra” extends to the following:

$$\begin{aligned} (dB_t^i)^2 &= dt, & dt \cdot dB_t^i &= 0, & i &= 1, 2; \\ dt \cdot dt &= 0, & dB_t^1 \cdot dB_t^2 &= 0. \end{aligned}$$

Hence,

$$\begin{aligned} (dX_t)^2 &= (\sigma_{11}^2 + \sigma_{12}^2)dt, \\ (dY_t)^2 &= (\sigma_{21}^2 + \sigma_{22}^2)dt, \\ dX_t \cdot dY_t &= (\sigma_{11}\sigma_{21} + \sigma_{12}\sigma_{22})dt; \end{aligned}$$

whereas $dt \cdot dX_t = dt \cdot dY_t = dt \cdot dt = 0$.

More generally, consider a differentiable function $f(t, x, y)$. As before, let $f'_t, f'_x, f'_y, f''_{xx}, f''_{yy}$ and f''_{xy} denote the partial derivatives. Then, Itô's formula takes the following form:

$$df(t, X_t, Y_t) = f'_t dt + f'_x dX_t + f'_y dY_t + \frac{1}{2} f''_{xx} (dX_t)^2 + \frac{1}{2} f''_{yy} (dY_t)^2 + f''_{xy} dX_t dY_t. \quad (46)$$

Note, the last term above is equal to

$$\frac{1}{2} f''_{xy} dX_t dY_t + \frac{1}{2} f''_{yx} dY_t dX_t,$$

as $f''_{xy} = f''_{yx}$.

As a special case, consider $f(t, x, y) = xy$. Then, the formula in (46) reduces to what's known as *Itô's product rule*:

$$d(X_t Y_t) = X_t dY_t + Y_t dX_t + dX_t dY_t \quad (47)$$

7.2 Self-Financing Trading

Recall V_t in §6, the capital (or portfolio) that was used to carry out the delta hedging. Our starting point was dV_t in (40); and, in fact, we never needed V_t , (Note the terms involving V_t were cancelled in (42)). However, we can express V_t as follows:

$$V_t = \Delta_t S_t + \gamma_t m_t, \quad (48)$$

where γ_t is the number of units invested in the money-market account, in which one dollar at time zero becomes $m_t := e^{rt}$ dollars at t . (Clearly, γ_t is a random process.)

Hence, applying Itô's product rule, we have³

$$\begin{aligned} dV_t &= \Delta_t dS_t + S_t d\Delta_t + dS_t \cdot d\Delta_t \\ &\quad + \gamma_t dm_t + m_t d\gamma_t + dm_t \cdot d\gamma_t. \end{aligned}$$

Comparing the above with (40), where $V_t - \Delta_t S_t = \gamma_t m_t$ following (48), we have

$$S_t d\Delta_t + dS_t \cdot d\Delta_t + m_t d\gamma_t + dm_t \cdot d\gamma_t = 0. \quad (49)$$

The above equation can be viewed as a *self-financing condition*, (which $d\Delta_t$ and $d\gamma_t$ must satisfy).

To better understand the self-financing condition, consider its discrete-time counterpart:

$$\begin{aligned} S_n(\Delta_{n+1} - \Delta_n) + (S_{n+1} - S_n)(\Delta_{n+1} - \Delta_n) \\ + m_n(\gamma_{n+1} - \gamma_n) + (m_{n+1} - m_n)(\gamma_{n+1} - \gamma_n) = 0, \end{aligned}$$

³Here, and below in (56), the term $dm_t \cdot d\gamma_t = re^{rt} dt \cdot d\gamma_t$ can be ignored. In general, however, if the value of money is replaced by a random process, then this term should be included.

where $m_n := (1 + r)^n$ is the discrete equivalent of e^{rt} . The above simplifies to

$$S_{n+1}(\Delta_{n+1} - \Delta_n) + m_{n+1}(\gamma_{n+1} - \gamma_n) = 0. \quad (50)$$

Note the two terms on the LHS correspond to rebalancing (from period n to period $n + 1$) the positions in stock and in cash, respectively. Since the two terms sum to zero, money must be taken from one position and put into the other. In this sense, the hedging portfolio is “self-financing”, i.e., no extra cash beyond the initial capital is required to fund the hedging strategy.

Indeed, the discrete-time self-financing condition in (50) can be directly derived as follows (which is quite similar to the derivation in the continuous setting). In period n , we have

$$V_n = \Delta_n S_n + \gamma_n m_n. \quad (51)$$

The gain from period n to period $n + 1$ is:

$$V_{n+1} - V_n = \Delta_n(S_{n+1} - S_n) + \gamma_n(m_{n+1} - m_n),$$

i.e., the sum of the gain in the stock position and the income (interest) from cash. Hence, substituting (51) into the above, we have:

$$V_{n+1} = \Delta_n S_{n+1} + \gamma_n m_{n+1}. \quad (52)$$

But, like (51), we must also have

$$V_{n+1} = \Delta_{n+1} S_{n+1} + \gamma_{n+1} m_{n+1}. \quad (53)$$

Equating the RHS's of (52) and (53) yields the condition in (50).

7.3 A Corrected Derivation of the PDE

A common criticism of the derivation of the PDE in §4 is that the $d\Pi_t$ expression in (26) (and in (27) as well) has failed to apply Itô's product rule to the second term of Π_t , namely, $\Delta_t S_t$. When that is corrected, two more terms will appear on the right hand side of (27):

$$d\Pi_t = -\left(\frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 f}{\partial S_t^2} + \frac{\partial f}{\partial t}\right)dt + S_t d\Delta_t + dS_t \cdot d\Delta_t. \quad (54)$$

Recall, the portfolio Π_t consists of $\Delta_t := \frac{\partial f}{\partial S_t}$ shares of the underlying stock and -1 share of the option. Hence,

$$\Pi_t = \Delta_t S_t - f(t, S_t) = \Delta_t S_t - V_t = -\gamma_t m_t, \quad (55)$$

where the second equation is due to $V_t = f(t, S_t)$ in the delta hedging strategy, and the third equation follows from (48). Applying Itô's product rule to (55), we have

$$d\Pi_t = -\gamma_t dm_t - m_t d\gamma_t - dm_t \cdot d\gamma_t. \quad (56)$$

Equating the RHS of the above with the RHS of (54) and taking into account the self-financing condition in (49), we have

$$-\left(\frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 f}{\partial S_t^2} + \frac{\partial f}{\partial t}\right)dt = -\gamma_t dm_t = \Pi r dt,$$

where the last equation follows from (55). (Recall, $m_t := e^{rt}$, hence, $dm_t = rm_t dt$.) In other words, the PDE in (29), which follows from (28), is, after all, correct, even though the derivation there wasn't quite right!

Exercise. Consider the two-dimensional BM model in §7.1.

(i) Derive $\text{Cov}(dX_t, dY_t)$.

(ii) If we write

$$dX_t = \mu_1 dt + \sigma_X dW_t^1, \quad dY_t = \mu_2 dt + \sigma_Y dW_t^2,$$

with W_t^1 and W_t^2 being two BM's; what is σ_X and σ_Y , and what is the correlation between W_t^1 and W_t^2 ?

(iii) Extend (46) and (47) to the general multi-dimensional case, i.e., when there are ℓ independent BM's, and m processes:

$$X_t^i = \mu_i t + \sigma_{i1} B_t^1 + \sigma_{i2} B_t^2 + \cdots + \sigma_{i\ell} B_t^\ell, \quad i = 1, \dots, m.$$

Exercise. The stock price S_t follows GBM, $dS_t = S_t(\mu dt + \sigma dB_t)$ as before. Define another process,

$$\zeta_t := \exp\left[-\left(r + \frac{1}{2}\lambda^2\right)t - \lambda B_t\right],$$

where $\lambda := (\mu - r)/\sigma$, and r is the risk-free interest rate. (i) Verify that

$$d\zeta_t = -\zeta_t[r dt + \lambda dB_t].$$

Let V_t denote the value of a portfolio that invests in Δ_t shares of the stock and the rest in cash, just like the hedging portfolio in §6 but Δ_t need not be the option sensitivity. (ii) Verify that

$$dV_t = rV_t dt + \Delta_t(\mu - r)S_t dt + \Delta_t \sigma S_t dB_t.$$

(iii) Derive $d(\zeta_t V_t)$ using Itô's product rule.

Appendix: Quadratic Variation and Total Variation of BM

Below we shall use both B_t and $B(t)$ denote the standard Brownian motion – a more elaborate time index will be denoted as an argument (instead of a subscript).

There are two types of convergence in the theorem below, a.s. (almost surely, i.e., probability one) and L_2 (in the sense of L_2 norm, or the second moment). The B-C lemma refer to the Borel-Cantelli Lemma; refer to, for instance, S.M. Ross, *Stochastic Processes*, Wiley, 1996 (pp. 4-6).

Fix a time $t > 0$, and divide the interval $[0, t]$ into 2^n equal segments. Denote

$$\Delta_k(n) := B\left(\frac{k}{2^n}t\right) - B\left(\frac{k-1}{2^n}t\right).$$

Then, $\Delta_k(n)$, for $k = 1, \dots, 2^n$, are iid $N(0, t/2^n)$ random variables.

The total variation (TV) and quadratic variation (QV) of the BM are defined as follows:

$$(\text{TV}) := \lim_{n \rightarrow \infty} \sum_{k=1}^{2^n} |\Delta_k(n)|, \quad (\text{QV}) := \lim_{n \rightarrow \infty} \sum_{k=1}^{2^n} [\Delta_k(n)]^2. \quad (57)$$

Theorem 1 (i) $(\text{QV}) = t$, in both a.s. and L_2 . (ii) $(\text{TV}) = \infty$.

Proof. First we show (ii) is implied by (i). Let

$$\Delta_\ell(n) := \max_{1 \leq k \leq 2^n} \{\Delta_k(n)\}.$$

To simplify notation, omit the argument n . We have

$$|\Delta_\ell| \sum_{k=1}^{2^n} |\Delta_k| \geq \sum_{k=1}^{2^n} \Delta_k^2(n);$$

or

$$\sum_{k=1}^{2^n} |\Delta_k| \geq \frac{\sum_{k=1}^{2^n} \Delta_k^2(n)}{|\Delta_\ell|}.$$

When $n \rightarrow \infty$, the path continuity of BM implies $\Delta_\ell \rightarrow 0$. Hence, the right hand side above goes to infinity, as the numerator goes to a finite t , following (i).

To prove (i), write the result as

$$\sum_{k=1}^{2^n} \left(\Delta_k^2 - \frac{t}{2^n}\right) \rightarrow 0. \quad (58)$$

First, we show this holds in the L_2 sense, i.e.,

$$\mathbb{E} \left[\sum_{k=1}^{2^n} \left(\Delta_k^2 - \frac{t}{2^n}\right) \right]^2 \rightarrow 0. \quad (59)$$

Note that the cross terms on the left hand side above all vanish due to independent increments; i.e., for all $j \neq k$,

$$\mathbb{E}[(\Delta_j^2 - \frac{t}{2^n})(\Delta_k^2 - \frac{t}{2^n})] = \mathbb{E}(\Delta_j^2 - \frac{t}{2^n}) \cdot \mathbb{E}(\Delta_k^2 - \frac{t}{2^n}) = 0.$$

Hence, the left hand side of (59) is equal to

$$\mathbb{E}\left[\sum_{k=1}^{2^n}(\Delta_k^2 - \frac{t}{2^n})\right]^2 = \sum_{k=1}^{2^n} \mathbb{E}(\Delta_k^2 - \frac{t}{2^n})^2.$$

We have, with Z denoting the standard normal variate,

$$\mathbb{E}(\Delta_k^2 - \frac{t}{2^n})^2 = \text{Var}(\Delta_k^2) = \text{Var}(\frac{t}{2^n} Z^2) = \frac{t^2}{2^{2n}} [\mathbb{E}(Z^4) - \mathbb{E}^2(Z^2)] = \frac{2t^2}{2^{2n}}.$$

(Note that $\mathbb{E}(Z^4) = 3$.) Hence,

$$\mathbb{E}\left[\sum_{k=1}^{2^n}(\Delta_k^2 - \frac{t}{2^n})\right]^2 = 2^n \frac{2t^2}{2^{2n}} = \frac{2t^2}{2^n} \rightarrow 0.$$

To show the a.s. convergence in (58), making use of the above, along with Chebyshev's inequality, we have

$$\mathbb{P}\left[\left|\sum_{k=1}^{2^n}(\Delta_k^2 - \frac{t}{2^n})\right| > \epsilon\right] \leq \frac{1}{\epsilon^2} \cdot \frac{2t^2}{2^n}. \quad (60)$$

The right hand side above being a summable sequence, a.s. convergence follows from the B-C lemma. $\square$

Remark 2 (i) Since the Brownian path is continuous over the interval $[0, t]$, $(TV) = \infty$ means that the path must have an infinite number of ups and downs, each of which, however, is infinitesimal, taking into account $(QV) = t$. This points to the extreme zig-zagness of the path, and hence, its non-differentiability.

(ii) The limit $(QV) = t$ can be expressed as follows:

$$\int_0^t [dB_s]^2 = t.$$

Since $t = \int_0^t ds$, we can write

$$[dB_s]^2 = ds,$$

which is the relation in (1).