Supplementary Notes on "Mathematical Methods in Finance"

Chenxu Li *
Department of Business Statistics
and Econometrics
Guanghua School of Management
Peking University

Saturday 3rd October, 2015

Abstract

In this set of notes, we provide some technical details supplementary to our lecture slides from Lecture 5 to Lecture 10. As this is a very preliminary version, please keep them for your own use and don't circulate them. We will try to write them into a text book. I appreciate any reports of errors and typos. With this set of notes, you don't need worry about the problem of viewing clustered formula derivations the board from a distant point in our big classroom. However, please note that it is not an exhaustive list of our in-class discussion. A combination of course slides, these supplementary notes, homework problem sets and your notes taken from our in-class discussions would be quite helpful to your study.

1 Lecture 5

1.1 Variations of a continuously differentiable function

Let $f(t)$ be a second-order differentiable function defined on $0 \leq t \leq T$. The first-order variation of f up to time T is defined as

$$FV_f(T) := \lim_{||\Pi|| \to 0} \sum_{j=0}^{n-1} |f(t_{j+1}) - f(t_j)|,$$

where $\Pi = \{t_0, t_1, \cdots, t_n\}$ is a partition with $0 = t_0 < t_1 < \cdots < t_n = T$ and

$$||\Pi|| = \max\{t_{j+1} - t_j,\ j = 0, 1, 2, ..., n-1\}.$$

*Email: cxli@gsm.pku.edu.cn

Then, we have

$$\lim_{||\Pi||\to 0}\sum_{j=0}^{n-1}|f(t_{j+1})-f(t_j)| = \lim_{||\Pi||\to 0}\sum_{j=0}^{n-1}|f'(t_j^*)(t_{j+1}-t_j)| = \int_0^T |f\prime(t)|dt < +\infty.$$

The first-order variation of a continuously differentiable function is finite.

The quadratic variation of f up to time T is defined as

$$QV_f(T) := \lim_{||\Pi||\to 0}\sum_{j=0}^{n-1}|f(t_{j+1})-f(t_j)|^2,$$

We note that

$$\sum_{j=0}^{n-1}[f(t_{j+1})-f(t_j)]^2 = \sum_{j=0}^{n-1}|f\prime(t_j^*)|^2(t_{j+1}-t_j)^2 \le ||\Pi||\sum_{j=0}^{n-1}|f\prime(t_j^*)|^2(t_{j+1}-t_j).$$

Thus, we have

$$\begin{aligned}
QV_f(T) &\le \lim_{||\Pi||\to 0}\left[||\Pi||\sum_{j=0}^{n-1}|f\prime(t_j^*)|^2(t_{j+1}-t_j)\right] \\
&= \lim_{||\Pi||\to 0}||\Pi|| \lim_{||\Pi||\to 0}\sum_{j=0}^{n-1}|f\prime(t_j^*)|^2(t_{j+1}-t_j) \\
&= \lim_{||\Pi||\to 0}||\Pi||\cdot\int_0^T |f\prime(t)|^2 dt = 0.
\end{aligned}$$

1.2 Variations of a Brownian Motion

Denote by

$$Q_\Pi := \sum_{j=0}^{n-1}|W(t_{j+1})-W(t_j)|^2$$

The quadratic variation is thus

$$QV_W(T) = \lim_{||\Pi||\to 0}Q_\Pi = \lim_{||\Pi||\to 0}\sum_{j=0}^{n-1}|W(t_{j+1})-W(t_j)|^2.$$

We note that Q_Π is a sum of independent random variables. Therefore, its mean and variance are the sums of the means and variances of these random variables, respectively. Because

$$\mathbb{E}\left(W(t_{j+1}) - W(t_j)\right)^2 = \mathtt{var}[W(t_{j+1}) - W(t_j)] = t_{j+1} - t_j,$$

we have

$$\mathbb{E}(Q_\Pi) \equiv \sum_{j=0}^{n-1} \mathbb{E}(W(t_{j+1}) - W(t_j))^2 = \sum_{j=0}^{n-1} (t_{j+1} - t_j) \equiv T.$$

Moreover, we have

$$\begin{aligned} \text{var}\left[(W(t_{j+1}) - W(t_j))^2\right] &= \mathbb{E}\left[\left((W(t_{j+1}) - W(t_j))^2 - (t_{j+1} - t_j)\right)^2\right] \\ &= \mathbb{E}\left[(W(t_{j+1}) - W(t_j))^4\right] - 2(t_{j+1} - t_j)\mathbb{E}\left(W(t_{j+1}) - W(t_j)\right)^2 + (t_{j+1} - t_j)^2. \end{aligned}$$

It can be easily proved that, for a variable X following normal distribution $\mathcal{N}(0, \sigma^2)$, we have

$$\mathbb{E}\left(X^4\right) = 3\sigma^4.$$

Therefore, we have

$$\begin{aligned} \text{var}\left[(W(t_{j+1}) - W(t_j))^2\right] &= 3(t_{j+1} - t_j)^2 - 2(t_{j+1} - t_j)^2 + (t_{j+1} - t_j)^2 \\ &= 2(t_{j+1} - t_j)^2, \end{aligned}$$

Thus, we have

$$\begin{aligned}
\mathrm{var}(Q_\Pi) &= \mathrm{var}\left[\sum_{j=0}^{n-1}(W(t_{j+1})-W(t_j))^2\right] \\
&= \sum_{j=0}^{n-1}\mathrm{var}\left[(W(t_{j+1})-W(t_j))^2\right] \\
&= \sum_{j=0}^{n-1} 2(t_{j+1}-t_j)^2 \\
&\leq \sum_{j=0}^{n-1} 2||\Pi||(t_{j+1}-t_j) = 2||\Pi||T.
\end{aligned}$$

So, we have $\lim_{||\Pi||\to 0}\mathrm{var}(Q_\Pi) \equiv 0$. Hence, if follows that

$$\lim_{||\Pi||\to 0} Q_\Pi \equiv \mathbb{E}(Q_\Pi) = T. \tag{1.1}$$

The first-order variation is defined as

$$FV_W(T) = \lim_{||\Pi||\to 0}\sum_{j=0}^{n-1}|W(t_{j+1})-W(t_j)|.$$

It is obvious that

$$Q_\Pi \leq \max|W(t_{j+1})-W(t_j)|\cdot\sum_{j=0}^{n-1}|W(t_{j+1})-W(t_j)| \tag{1.2}$$

Suppose

$$\lim_{||\Pi||\to 0}\sum_{j=0}^{n-1}|W(t_{j+1})-W(t_j)| < +\infty.$$

By the continuity property of Brownian motion, we have

$$\lim_{||\Pi||\to 0}\max|W(t_{j+1})-W(t_j)| = 0.$$

Thus, by taking limit on the both sides of the inequality, we have

$$\lim_{||\Pi||\to 0} Q_\Pi \le \lim_{||\Pi||\to 0} \max |W(t_{j+1}) - W(t_j)| \cdot \lim_{||\Pi||\to 0} \sum_{j=0}^{n-1} |W(t_{j+1}) - W(t_j)| = 0.$$

This implies that

$$\lim_{||\Pi||\to 0} Q_\Pi = 0,$$

which contradicts to (1.1). Hence, we must have

$$FV_W(T) = \lim_{||\Pi||\to 0} \sum_{j=0}^{n-1} |W(t_{j+1}) - W(t_j)| = +\infty.$$

1.3 On multidimensional Brownian motions

Suppose $\{(B_1(t), B_2(t), \cdots, B_d(t))\}$ is a d dimensional Brownian motion with covarian matrix $(t-s)\Sigma$, where $\Sigma = (\rho_{ij})_{d\times d}$. Denote by

$$B(t) = (B_1(t), B_2(t), \cdots, B_d(t))^\top,$$

a column random vector. In this case, we knew that

$$t\Sigma = \text{cov}(B(t), B(t)) \equiv (\text{cov}(B_i(t), B_j(t)))_{d\times d}.$$

Thus, we have the following elementwise form

$$t\rho_{ij} = \text{cov}(B_i(t), B_j(t)).$$

We show that ρ_{ij} is the correlation between $B_i(t)$ and $B_j(t)$. Because var$B_i(t)$ =var$B_j(t) = t$, it is easy to see that

$$\text{corr}(B_i(t), B_j(t)) = \frac{\text{cov}(B_i(t), B_j(t))}{\sqrt{\text{var}(B_i(t))}\sqrt{\text{var}(B_j(t))}} = \frac{\text{cov}(B_i(t), B_j(t))}{\sqrt{t}\sqrt{t}} = \rho_{ij}.$$

That is to say ρ_{ij} represents the correlation between $B_i(t)$ and $B_j(t)$.

Suppose we can find matrix A such that

$$AA^\top = \Sigma. \tag{1.3}$$

Then, we can construct $B(t)$ by $B(t) = AZ(t)$, where $\{Z(t)\}$ is a standard d dimensional Brownian motion. Indeed, we just need to check all the points in the definition are satisfied. In particular, to check the covariance matrix, we deduce that

$$\begin{aligned}
\text{cov}(AZ(t), AZ(t)) &= \mathbb{E}[AZ(t)(AZ(t))^\top] - \mathbb{E}(AZ(t))\mathbb{E}(AZ(t))^\top] \\
&= \mathbb{E}[AZ(t)(AZ(t))^\top] \\
&= \mathbb{E}[AZ(t)Z(t)^\top A^\top] \\
&= A\mathbb{E}[Z(t)Z(t)^\top]A^\top \\
&= AIA^\top = AA^\top = \Sigma.
\end{aligned}$$

This can be written in elementwise form (to add in the book.) In linear algebra, (1.3) is called Cholesky decomposition. In particuar the matrix A can be chosen as a lower diagonal matrix.

2 Lecture 6

2.1 Properties of Itô integral $I(t)$ for simple processes $\triangle(t)$

2.1.1 $I(t)$ is continuous in t

Though $I(t)$ is not continuous if we focus on grid points $\{t_k\}$, we have the continuity of $I(t)$ in t. A simple reason is as follows. For any arbitrary $t^* \in [t_k, t_{k+1})$, we have

$$\begin{aligned}
\lim_{t \to t^*} I(t) &= \lim_{t \to t^*} \left(\sum_{j=0}^{k-1} \Delta(t_j)[W(t_{j+1}) - W(t_j)] + \Delta(t_k)[W(t) - W(t_k)] \right) \\
&= I(t_k) + \lim_{t \to t^*} \Delta(t_k)[W(t) - W(t_k)] \\
&= I(t_k) + I(t^*) - I(t_k) \\
&= I(t^*).
\end{aligned}$$

2.1.2 $I(t)$ is $\mathcal{F}(t)$-measurable

For any $t \in [t_k, t_{k+1})$,

$$I(t) = \sum_{j=0}^{k-1} \Delta(t_j)[W(t_{j+1}) - W(t_j)] + \Delta(t_k)[W(t) - W(t_k)].$$

We note that $\Delta(t_0), \Delta(t_1), ...,$ and $\Delta(t_k)$ are $\mathcal{F}(t)$-measurable; $W(t_0), W(t_1), ...,$ and $W(t)$ are $\mathcal{F}(t)$-measurable. Thus, $I(t)$ is $\mathcal{F}(t)$-measurable;

2.1.3 $I(t)$ is a martingale

For any $t > s$, we will prove that

$$\mathbb{E}\left[I(t)|\mathcal{F}(s)\right] = I(s). \tag{2.1}$$

Indeed, we can easily work with the grid points as follows:

$$\begin{aligned}
\mathbb{E}\left[I(t_{k+1})|\mathcal{F}(t_k)\right] &= \mathbb{E}\left(\sum_{j=0}^{k} \Delta(t_j)[W(t_{j+1}) - W(t_j)]|\mathcal{F}(t_k)\right) \\
&= I(t_k) + \mathbb{E}[\Delta(t_k)[W(t_{k+1}) - W(t_k)]|\mathcal{F}(t_k)] \\
&= I(t_k) + \Delta(t_k)\mathbb{E}[W(t_{k+1}) - W(t_k)|\mathcal{F}(t_k)] \\
&= I(t_k) + \Delta(t_k)\mathbb{E}[W(t_{k+1}) - W(t_k)] \\
&= I(t_k).
\end{aligned}$$

Similarly, by taking care of some detailed tricks, we can prove the martingale property (2.1) outside of grid points.

2.1.4 Itô Isometry

We will show that

$$\mathbb{E}I^2(t) = \mathbb{E}\left(\int_0^t \Delta^2(u)du\right).$$

Without loss of generality, we still prove it for the case when t is a grid point, say $t = t_k$. For simplicity, we denote by $W(t_{j+1}) - W(t_j) = a_j$. Thus, we have

$$I(t) = \sum_{j=0}^{k} \Delta(t_j) a_j$$

and

$$I^2(t) = \sum_{j=0}^{k} \Delta^2(t_j) {a_j}^2 + 2 \sum_{0 \le i \le j \le k} \Delta(t_i)\Delta(t_j) a_i a_j.$$

First, let's prove the expectations of cross terms are 0. Indeed, for each pair $i \le j$, we note that $\Delta(t_i)\Delta(t_j)a_i$ is $\mathcal{F}(t_j)$-measurable, a_j is independent of $\mathcal{F}(t_j)$, and $\mathbb{E}a_j = 0$. Therefore,

$$\begin{aligned}
\mathbb{E}[\Delta(t_i)\Delta(t_j)a_i a_j] &= \mathbb{E}[\mathbb{E}[\Delta(t_i)\Delta(t_j)a_i a_j | \mathcal{F}(t_j)]] \\
&= \mathbb{E}[\Delta(t_i)\Delta(t_j)a_i \mathbb{E}[a_j | \mathcal{F}(t_j)]] \\
&= \mathbb{E}[\Delta(t_i)\Delta(t_j)a_i \mathbb{E}[a_j]] \\
&= 0.
\end{aligned}$$

Then, we consider the squared terms. We have

$$\begin{aligned}
\mathbb{E}[\Delta^2(t_j){a_j}^2] &= \mathbb{E}[\mathbb{E}\left[\Delta(t_j)^2 {a_j}^2 | \mathcal{F}(t_j)\right]] \\
&= \mathbb{E}[\Delta(t_j)^2 \mathbb{E}[{a_j}^2 | \mathcal{F}(t_j)] \\
&= \mathbb{E}[\Delta(t_j)^2 (t_{j+1} - t_j)].
\end{aligned}$$

Thus, it is easy to have

$$\mathbb{E}I^2(t) = \sum_{j=0}^{k} \mathbb{E}[\Delta(t_j)^2 (t_{j+1} - t_j)] = \mathbb{E}\left(\int_0^t \Delta^2(u) du\right).$$

2.1.5 Quadratic Variation

We will show that

$$[I, I](t) = \int_0^t \Delta^2(u)du.$$

Without loss of generality, we still prove it for the case when t is a grid point, say $t = t_k$. Recall that quadratic variation of I is given by

$$\lim_{\|\Pi\| \to 0} \sum_{i=0}^{m-1} [I(\xi_{i+1}) - I(\xi_i)]^2,$$

where $\Pi = \{\xi_0, \xi_1, \xi_2, , ..., \xi_N\}$ gives a partition of the interval $[0, t]$ even finer than $\{t_0, t_1, t_2, , ..., t_n\}$. We consider partitions of the interval $[t_j, t_{j+1}]$ by taking $t_j = s_0 < s_1 < ... < s_m = t_{j+1}$. Because $\Delta(t_j)$ is constant on $[t_j, t_{j+1})$, we have

$$\sum_{i=0}^{m-1} [I(s_{i+1}) - I(s_i)]^2 = \Delta^2(t_j) \sum_{i=0}^{m-1} [W(s_{i+1}) - W(s_i)]^2.$$

Using the quadratic variation of Brownian motion, we obtain that

$$\lim_{\max\{s_{i+1} - s_i, i=0,1,2,...,m-1\} \to 0} \sum_{i=0}^{m-1} [W(s_{i+1}) - W(s_i)]^2 = t_{j+1} - t_j.$$

Thus, we have

$$\sum_{i=0}^{m-1} [I(s_{i+1}) - I(s_i)]^2 \to \Delta^2(t_j)(t_{j+1} - t_j) = \int_{t_j}^{t_{j+1}} \Delta^2(u)du \tag{2.2}$$

Summing up for all the intervals, by the definition of quadratic variation, we obtain that

$$[I, I](t) = \int_0^t \Delta^2(u)du.$$

2.2 Simple applications of the Itô formula

Let $f(x) = \frac{x^2}{2}$. By taking derivatives, we have $f'(x) = x$ and $f''(x) = 1$. Thus, by using the Itô formulas, we have

$$\begin{aligned} f(W(t)) &= \frac{1}{2}W^2(t) \\ &= \frac{1}{2}W^2(0) + \int_0^t W(u)dW(u) + \frac{1}{2}\int_0^t 1du \\ &= \int_0^t W(u)dW(u) + \frac{1}{2}t. \end{aligned}$$

This implies that

$$\int_0^t W(u)dW(u) = \frac{1}{2}(W^2(t) - t). \tag{2.3}$$

We have proved that, using the definition of Brownian motion previously, $\{W^2(t) - t\}$ is a martingale. Since the process of stochastic integrals $\{\int_0^t W(u)dW(u)\}$ is a martingale, the relation (2.3) renders an alternative proof of $\{W^2(t) - t\}$ being a martingale.

We suggest readers to consider the following exercise

$$\int_0^t W(u)^n dW(u)$$

for any arbitrary integer n. And further think about the following popular interview question: can you use the Itô formula to find

$$\mathbb{E}W(t)^n = ?$$

Indeed, it is easy to set up an iteration relation. Please solve it!

2.3 Itô processes

Itô processes provide a broad ground for modeling asset returns. In particular, in recent years, the analysis of high-frequency trading date heavily hinges on the analysis and estimation of their quadratic variation. For the Itô process X, we have

$$[X,X](t)=\int_0^t \Delta^2(u)du. \tag{2.4}$$

The proof of this result can be seen in Lemma 4.4.4 in Shreve's book. It follows the definition of quadratic variation, i.e., to show that

$$\sum(X(\xi_{i+1})-X(\xi_i))^2 \to \int_0^t \Delta^2(u)du \tag{2.5}$$

in probability as the partition becomes finer and finer. Such a result reveals that the dt integrals don't contribute anything to quadratic variation. i.e.,

$$[X,X](t)=[I,I]\,(t)=\int_0^t \Delta^2(u)du, \tag{2.6}$$

where

$$I(t)=\int_0^t \Delta(u)dW(u). \tag{2.7}$$

We note that, when $\Delta \equiv 1$, (2.6) reduces to

$$[I,I]\,(t)=\int_0^t 1du=t.$$

In this case,

$$I(t)=\int_0^t \Delta(u)dW(u)=\int_0^t 1dW(u)=W(t).$$

Thus, if we view the stochastic integral (2.7) as a generalization of Brownian motion, (2.6) is a generalization of the quadratic variation property of Brownian motion, i.e., $[W,W](t)=t$. We also note that the quadratic variation (2.6) is stochastic in general, not necessarily to be a constant as in the case of Brownian motion. This can be easily seen from (2.5).

2.4 A useful generalization (with some bonus knowledge)

Before closing this section, we provide a further generalization. It is natural to consider stochastic integrals w.r.t. a general martingale. Suppose M is a continuous-time martingale, we can define

$$I^M(t) = \int_0^t \Delta(u) dM(u). \tag{2.8}$$

The way to define it follows the same spirit for defining stochastic integrals w.r.t. Browian motions. In this case, all the previous (measurability, linearity, martingale) properties hold in the same way except that the Itô isometry becomes

$$\mathbb{E} I^M(t)^2 = \mathbb{E}\left(\int_0^t \Delta^2(u) d[M, M](u)\right), \tag{2.9}$$

and quadratic variation becomes

$$[I^M, I^M](t) = \int_0^t \Delta^2(u) d[M, M](u). \tag{2.10}$$

Interested readers may wonder why Itô isometry and qudratic variation share the same elements in their expressions.

Indeed, via technical stochastic analysis and under some technical conditions, one is able to show that for any martingale $\{\mathcal{M}(t)\}$, the process $\{\mathcal{M}(t)^2 - [\mathcal{M}, \mathcal{M}](t)\}$ is a martingale. And, $A(t) = [\mathcal{M}, \mathcal{M}](t)$ is the only process making $\mathcal{M}(t)^2 - A(t)$ a martingale. This is obviously out of our reach in this course, which provides mathematical tools for research in finance and economics. Lucikly that we have seen special cases of this general theorem. For example, when $\mathcal{M}(t) = W(t)$, we have a martingale $\{W(t)^2 - [W, W](t)\} = \{W(t)^2 - t\}$ is a martingale. In addition, for the martingale

$$\mathcal{M}(t) = I(t) = \int_0^t \Delta(u) dW(u),$$

we will have a martingale constitued by

$$\mathcal{M}(t)^2 - [\mathcal{M}, \mathcal{M}](t) = I(t)^2 - [I, I](t).$$

Thus, if one has the quadatic variation property

$$[I, I](t) = \int_0^t \Delta^2(u) du,$$

by the martingale property, it is easy to have

$$\mathbb{E}[\mathcal{M}(t)^2 - [\mathcal{M}, \mathcal{M}](t)] = \mathbb{E}[I(t)^2 - [I, I](t)] = 0,$$

i.e.,

$$\mathbb{E}I(t)^2 = \mathbb{E}[I, I](t) = \mathbb{E}\left(\int_0^t \Delta^2(u) du\right).$$

Similarly, for the case of (2.8), the aforementioned general property implies that (2.10) leads to (2.9). Thus, the quadratic variation (2.10) plays an important role; it can be heuristically understood easily from the following chart:

$$\begin{array}{ll} \sum & (I(\xi_{i+1}) - I(\xi_i))^2 \\ \downarrow & \qquad\downarrow \\ \int_0^t & \Delta^2(u) d[M, M](u) \end{array}.$$

Here, the term $\Delta^2(u)d[M, M](u)$ is an infinitesimal analogy of (2.2).

2.5 Multivariate stochastic calculus

2.5.1 Definition of covariation

We first introduce the notion of covariation as a natural generalization of quadratic variation. Suppose we have two stochastic processes $\{X(t)\}$ and $\{Y(t)\}$. The covariation between the two can be defined a stochastic process satisfying

$$[X,Y](t) = \lim_{\|\Pi\|\to 0} \sum_{k=0}^{n-1} (X(t_{k+1}) - X(t_k))(Y(t_{k+1}) - Y(t_k)), \tag{2.11}$$

where we recall that $\Pi = \{t_0, t_1, t_2, ..., t_n\}$ is a partition of the interval $[0,t]$ with $t_0 = 0$ and $t_n = t$. Here, the normal $\|\Pi\| = \max\{t_j - t_{j-1}, j = 1, 2, ..., n\}$. We note that the lim in (2.11) is in the sense of probability. When $X = Y$, we have $[X,Y](t) = [X,X](t)$ is the quadratic variation of X. We mention the following useful properties without proofs, which naturally follows from tedious and straightforward applications of the definition (2.11):

1. symmetric property: $[X,Y](t) = [Y,X](t)$;
2. linearity: $[X, Y+Z](t) = [X,Y](t) + [X,Z](t)$ and $[X, cY](t) = c[X,Y](t)$.

2.5.2 On multidimensional Brownian motions

From the slides, we have

$$d[W_i, W_i](t) \equiv dW_i(t)dW_i(t) = dt,$$

and

$$d[W_i, W_j](t) \equiv dW_i(t)dW_j(t) = 0, \text{ for } i \neq j. \tag{2.12}$$

In addition, we recall that

$$d[W_i, \cdot](t) \equiv dW_i(t)dt = 0.$$

Here, to show (2.12), it is equivalent to prove

$$[W_i, W_j](t) = \lim_{\|\Pi\|\to 0} \sum_{k=0}^{n-1} (W_i(t_{k+1}) - W_i(t_k))(W_j(t_{k+1}) - W_j(t_k)) = 0.$$

which can be regarded as an excellent exercise. How? Find it in Shreve?

Here is an interesting exercise enhancing our understanding of multidimensional Brownian motions. Suppose $\{(B_1(t), B_2(t), \cdots, B_d(t))\}$ is a d dimensional Brownian motion with covarian matrix $(t-s)\Sigma$, where $\Sigma = (\rho_{ij})_{d\times d}$. Denote by

$$B(t) = (B_1(t), B_2(t), \cdots, B_d(t))^\top$$

a column random vector. In this case, we knew that

$$t\Sigma = \mathrm{cov}(B(t), B(t)) \equiv (\mathrm{cov}(B_i(t), B_j(t)))_{d\times d}.$$

Thus, we have the following elementwise form

$$t\rho_{ij} = \mathrm{cov}(B_i(t), B_j(t)) = \mathbb{E}[B_i(t)B_j(t)].$$

So, it is easy to see that

$$\rho_{ij} = \frac{\mathrm{cov}(B_i(t), B_j(t))}{\sqrt{\mathrm{var}(B_i(t))}\sqrt{\mathrm{var}(B_j(t))}} = \frac{\mathbb{E}[B_i(t)B_j(t)]}{\sqrt{\mathrm{var}(B_i(t))}\sqrt{\mathrm{var}(B_j(t))}} = \mathrm{corr}(B_i(t), B_j(t)).$$

That is to say ρ_{ij} represents the correlation between $B_i(t)$ and $B_j(t)$. Now, we claim that ρ_{ij} renders the covariation between $B_i(t)$ and $B_j(t)$.

A simple proof is given as follows. By the Cholesky decomposition, we have the following construction of the two-dimensional Brownian motion $\{(B_i(t), B_j(t))\}$, i.e.,

$$\begin{aligned} B_i(t) &= Z_i(t), \\ B_j(t) &= \rho_{ij}Z_i(t) + \sqrt{1-\rho_{ij}^2}Z_j(t), \end{aligned}$$

where Z_i and Z_j are two independent standard one-dimensional Brownian motions. Thus, we have

$$\begin{aligned} [B_i, B_j](t) &= [Z_i, \rho_{ij}Z_i + \sqrt{1-\rho_{ij}^2}Z_j](t) \\ &= \rho_{ij}[Z_i, Z_i](t) + \sqrt{1-\rho_{ij}^2}[Z_i, Z_j](t) = \rho_{ij}t, \end{aligned}$$

where the second equality follows from the linearity property of covariation. Equivalent notation is

$$d[B_i, B_j](t) = dB_i(t)dB_j(t) = \rho_{ij}dt.$$

2.5.3 A quick path to multivariate stochastic calculus

We start from the two-dimensional case. Suppose we have two Itô processes defined by

$$\begin{aligned} X(t) &= X(0) + \int_0^t \Theta_1(u)du + \int_0^t \sigma_{11}(u)dW_1(u) + \int_0^t \sigma_{12}(u)dW_2(u), \\ Y(t) &= Y(0) + \int_0^t \Theta_2(u)du + \int_0^t \sigma_{21}(u)dW_1(u) + \int_0^t \sigma_{22}(u)dW_2(u). \end{aligned}$$

We begin by investigating their quadratic variation and covariation. Without proof, we provide the following useful property as a tool for deriving what we need here. Suppose we have the following two stochastic integrals w.r.t. general maringales (as defined in (2.8)):

$$\begin{aligned} I^M(t) &= \int_0^t \Delta_1(u)dM(u), \\ I^N(t) &= \int_0^t \Delta_2(u)dN(u), \end{aligned}$$

where M and N are two martingales. For example, we can choose M and N as two Brownian motions, e.g.,

$$M(t) = W_1(t) \text{ and } N(t) = W_2(t).$$

We have their covariation as

$$[I^M, I^N](t) = \int_0^t \Delta_1(u)\Delta_2(u)d[M,N](u). \tag{2.13}$$

In particualr, when $M = N$ and $\Delta_1 = \Delta_2$, this property reduces to (2.10); when $M(t) = W_1(t)$ and $N(t) = W_2(t)$, we have

$$[I^{W_1}, I^{W_2}](t) = \int_0^t \Delta_1(u)\Delta_2(u)d[W_1, W_2](u) = 0.$$

The principle behind it follows much the same way as what we have seen before. So, feel free to use it without proof. To summarise, the aforementioned properties will serve as a foundation for us to study multivariate stochastic calculus.

Indeed, similar to the case of (2.4), du integrals contribute nothing to convariations. Thus, we can use the linearity property of covariation to obtain that

$$\begin{aligned}
& [X,X](t) \\
= & \left[\int_0^{\cdot}\sigma_{11}(u)dW_1(u)+\int_0^{\cdot}\sigma_{12}(u)dW_2(u),\int_0^{\cdot}\sigma_{11}(u)dW_1(u)+\int_0^{\cdot}\sigma_{12}(u)dW_2(u)\right](t) \\
= & \left[\int_0^{\cdot}\sigma_{11}(u)dW_1(u),\int_0^{\cdot}\sigma_{11}(u)dW_1(u)\right](t)+\left[\int_0^{\cdot}\sigma_{11}(u)dW_1(u),\int_0^{\cdot}\sigma_{12}(u)dW_2(u)\right](t) \\
& +\left[\int_0^{\cdot}\sigma_{12}(u)dW_2(u),\int_0^{\cdot}\sigma_{11}(u)dW_1(u)\right](t)+\left[\int_0^{\cdot}\sigma_{12}(u)dW_2(u),\int_0^{\cdot}\sigma_{12}(u)dW_2(u)\right](t),
\end{aligned}$$

where using the covariation property (2.13) we obtain that

$$\begin{aligned}
\left[\int_0^{\cdot}\sigma_{11}(u)dW_1(u),\int_0^{\cdot}\sigma_{11}(u)dW_1(u)\right](t) &= \int_0^t\sigma_{11}(u)\sigma_{11}(u)d[W_1,W_1](t)=\int_0^t\sigma_{11}(u)^2du, \\
\left[\int_0^{\cdot}\sigma_{11}(u)dW_1(u),\int_0^{\cdot}\sigma_{12}(u)dW_2(u)\right](t) &= \int_0^t\sigma_{11}(u)\sigma_{12}(u)d[W_1,W_2](t)=0, \\
\left[\int_0^{\cdot}\sigma_{12}(u)dW_2(u),\int_0^{\cdot}\sigma_{11}(u)dW_1(u)\right](t) &= \int_0^t\sigma_{12}(u)\sigma_{11}(u)d[W_1,W_2](t)=0, \\
\left[\int_0^{\cdot}\sigma_{12}(u)dW_2(u),\int_0^{\cdot}\sigma_{12}(u)dW_2(u)\right](t) &= \int_0^t\sigma_{12}(u)\sigma_{12}(u)d[W_2,W_2](t)=\int_0^t\sigma_{12}(u)^2du.
\end{aligned}$$

Thus, we have

$$[X,X](t)=\int_0^t\sigma_{11}(u)^2dt+\int_0^t\sigma_{12}(u)^2dt.$$

Similarly, we have

$$[Y,Y](t)=\int_0^t\sigma_{21}(u)^2dt+\int_0^t\sigma_{22}(u)^2dt,$$

and

$$[X,Y](t)=\int_0^t\sigma_{11}(u)\sigma_{21}(u)dt+\int_0^t\sigma_{12}(u)\sigma_{22}(u)dt.$$

The above latter two properties can be regarded as excellent exercises. Using differential notation, we get

$$\begin{aligned}
dX(t)dX(t) &= d[X,X](t) = [\sigma_{11}(t)^2 + \sigma_{12}(t)^2]dt, \\
dY(t)dY(t) &= d[Y,Y](t) = [\sigma_{21}(t)^2 + \sigma_{22}(t)^2]dt, \\
dX(t)dY(t) &= d[X,Y](t) = [\sigma_{11}(t)\sigma_{21}(t) + \sigma_{21}(t)\sigma_{22}(t)]dt.
\end{aligned}$$

Now, it is easy to follow the slides to discuss the two-dimensional Itô formula.

Here is an exercise for the multivariate Itô formula. Assuming we know

$$W_3(t) = \rho W_1(t) + \sqrt{1-\rho^2}W_2(t),$$

we can use the Itô product formula (a straightforward implication of the two-dimensional Itô formula) to show that $\mathrm{corr}(W_1(t), W_3(t)) = \rho$. Indeed, we have

$$dW_1(t)W_3(t) = W_1(t)dW_3(t) + W_3(t)dW_1(t) + d[W_1, W_3](t).$$

Thus, we have

$$\begin{aligned}
\mathbb{E}[W_1(t)W_3(t)] &= \mathbb{E}\left(\int_0^t W_1(s)dW_3(s) + \int_0^t W_3(s)dW_1(s) + [W_1, W_3](t)\right) \\
&= \mathbb{E}\left(0 + 0 + [W_1, W_3](t)\right) \\
&= \mathbb{E}[W_1, W_3](t),
\end{aligned}$$

where the second equality follows from the martingale property of stochastic integrals. We note that

$$[W_1, W_3](t) = [W_1, \rho W_1 + \sqrt{1-\rho^2}W_2](t) = \rho t.$$

It follows that

$$\mathbb{E}[W_1(t)W_3(t)] = \mathbb{E}[W_1, W_3](t) = \rho t.$$

Thus, it is easy to have $\mathrm{corr}(W_1(t), W_3(t)) = \rho$.

Finally, we can generalize the Itô formula to any arbitrary n dimension. Let

$$X(t) = (X_1(t), X_2(t), ..., X_n(t)).$$

Suppose that for $i = 1, 2, ..., n$

$$X_i(t) = X_i(0) + \int_0^t \Theta_i(u)du + \sum_{k=1}^{d} \int_0^t \sigma_{ik}(u)dW_k(u).$$

Following the similar manner, we obtain that

$$dX_i(t)dX_j(t) = d[X_i, X_j](t) = \sum_{k=1}^{d} \sigma_{ik}(t)\sigma_{jk}(t)dt. \tag{2.14}$$

This is an excellent exercise. Now, using (2.14) as input, we have the following Itô formula for n dimension: denote by

$$df(t, X(t)) = \frac{\partial f}{\partial t}(t, X(t))dt + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(t, X(t))dX_i(t) + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(t, X(t))d[X_i, X_j](t).$$

3 Lecture 7

3.1 Definition of Multidimensional SDEs

Similar to one-dimensional SDEs, we can define multidimensional SDEs. A multi-dimensional stochastic differential equation is governed by

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

where $X(t)$ is an m dimensional stochastic process

$$X(t) = (X_1(t), X_2(t), \cdots, X_m(t))^\top;$$

$\mu(t,x)$ is an m dimensional vector

$$\mu(t,x) = (\mu_1(t,x), \mu_2(t,x), \cdots, \mu_m(t,x))^\top;$$

$\sigma(t,x)$ is an $m \times d$ matrix

$$\sigma(t,x) = \begin{pmatrix} \sigma_{11}(t,x), \sigma_{12}(t,x), \ldots, \sigma_{1d}(t,x) \\ \sigma_{21}(t,x), \sigma_{22}(t,x), \ldots, \sigma_{2d}(t,x) \\ \cdots \\ \sigma_{m1}(t,x), \sigma_{m2}(t,x), \ldots, \sigma_{md}(t,x) \end{pmatrix};$$

$W(t)$ is a standard d dimensional Brownian motions with

$$W(t) = (W_1(t), W_2(t), \cdots, W_d(t))^\top.$$

3.2 Explicitly solving the generalized geometric Brownian motion SDE

For $S(t)$ satisfies

$$dS(t) = \alpha(t)S(t)dt + \sigma(t)S(t)dW(t), \tag{3.1}$$

where $\alpha(t)$ and $\sigma(t)$ are processes adapted to the filtration $\{\mathcal{F}(t)\}$ associated with the Brownian motion $\{W(t)\}$. We note that, for the function

$$f(x) = \log x,$$

we calculus to obtain that

$$f'(x) = \frac{1}{x} \text{ and } f''(x) = -\frac{1}{x^2}.$$

Thus, by applying Itô formula, we obtain that

$$\begin{aligned} d\log S(t) &= \frac{1}{S(t)}dS(t) - \frac{1}{2}\frac{1}{S^2(t)}d[S,S](t) \\ &= \frac{1}{S(t)}\left[\alpha(t)S(t)dt + \sigma(t)S(t)dW(t)\right] - \frac{1}{2}\frac{1}{S^2(t)}\sigma(t)^2S(t)^2dt \\ &= \left(\alpha(t) - \frac{1}{2}\sigma^2(t)\right)dt + \sigma(t)dW(t) \end{aligned}$$

Integration of both sides yields that

$$\log S(t) - \log S(0) = \int_0^t \left(\alpha(u) - \frac{1}{2}\sigma^2(u)\right) dS + \int_0^t \sigma(u)dW(u),$$

which can be written equivalently as

$$S(t) = S(0)\exp\left(\int_0^t (\alpha(u) - \frac{1}{2}\sigma^2(u))du + \int_0^t \sigma(u)dW(u)\right).$$

3.3 Compute $\mathbb{E}R(t)$ and Var$R(t)$ for the CIR model

The Cox-Ingersoll-Ross model is given by

$$dR(t) = (\alpha - \beta R(t))dt + \sigma\sqrt{R(t)}dW(t),\ R(0) = r_0, \tag{3.2}$$

where α, β, σ are constants. Although we cannot derive an explicit solution for (3.2), we can use the Itô formula and solutions to ODEs to compute $\mathbb{E}R(t)$ and $\mathrm{Var}R(t)$ for the CIR Model.

We solve $\mathbb{E}R(t)$ first. Integretion of the both sides of (3.2) yields

$$R(t) - R(0) = \int_0^t (\alpha - \beta R(s))ds + \int_0^t \sigma\sqrt{R(s)}dW(s).$$

Taking expectation of both sides yields

$$\mathbb{E}R(t) - R(0) = \mathbb{E}\left[\int_0^t (\alpha - \beta R(s))ds\right] + \mathbb{E}\left(\int_0^t \sigma\sqrt{R(s)}dW(s)\right) \tag{3.3}$$

Recalling that the Itô integral $\int_0^t \sigma\sqrt{R(s)}dW(s)$ is a martingale, we have

$$\mathbb{E}\left(\int_0^t \sigma\sqrt{R(s)}dW(s)\right) = 0.$$

Thus, we have

$$\mathbb{E}R(t) - R(0) = \mathbb{E}\left[\int_0^t (\alpha - \beta R(s))ds\right].$$

An differential form of this equation reads

$$\begin{aligned} d\mathbb{E}R(t) &= [\alpha - \beta\mathbb{E}R(t)]dt, \\ \mathbb{E}R(0) &= r_0. \end{aligned}$$

Denote by $y(t) = \mathbb{E}R(t)$. We obtain an ODE with initial condition:

$$\begin{aligned} y'(t) &= \alpha - \beta y(t), \\ y(0) &= r_0 \end{aligned} \tag{3.4}$$

Multiplying $e^{\beta t}$ on the both sides of (3.4), we have

$$e^{\beta t}y'(t) + e^{\beta t}\beta y(t) = \alpha e^{\beta t},$$

which is equivalent to

$$[e^{\beta t}y(t)]' = \alpha e^{\beta t}.$$

Integrating the both sides, we have

$$e^{\beta t}y(t) = \int_0^t e^{\beta s}\alpha ds + C,$$

for some constants C. By plugging in $t = 0$, we have

$$y(0) = 0 + C = r_0.$$

Thus, we have

$$y(t) \equiv \mathbb{E}R(t) = e^{-\beta t}\left(r_0 + \int_0^t e^{\beta s}\alpha ds\right) = e^{-\beta t}r_0 + \frac{\alpha}{\beta}(1 - e^{-\beta t}) \tag{3.5}$$

Now, we solve for Var$R(t)$. We beging by noticing that

$$\text{var}R(t) = \mathbb{E}R^2(t) - (\mathbb{E}R(t))^2$$

As we already obtained $\mathbb{E}R(t)$, we still need $\mathbb{E}R^2(t)$. Applying the Itô Formula, we have

$$\begin{aligned} dR^2(t) &\equiv 2R(t)dR(t) + \frac{1}{2} \times 2d[R,R](t) \\ &\equiv 2R(t)[(\alpha - \beta R(t))dt + \sigma\sqrt{R(t)}dW(t)] + (\sigma\sqrt{R(t)})^2 dt \end{aligned}$$

Integating the both sides and taking expectations yield that

$$\begin{aligned} &\mathbb{E}R^2(t) - R^2(0) \\ = \;& 2\mathbb{E}\left(\int_0^t R(u)(\alpha - \beta R(u))du\right) + 2\sigma\mathbb{E}\left(\int_0^t R(u)\sqrt{R(u)}dW(u)\right) + \mathbb{E}\left(\int_0^t (\sigma\sqrt{R(u)})^2 du\right) \\ = \;& 2\int_0^t \mathbb{E}\left[R(u)(\alpha - \beta R(u))\right] du + 2\sigma\mathbb{E}\left(\int_0^t R(u)\sqrt{R(u)}dW(u)\right) + \sigma^2\int_0^t \mathbb{E}R(u)du \end{aligned}$$

Because the Itô integral $\int_0^t \sigma R(u)\sqrt{R(u)}dW(u)$ is a martingale, we have

$$\mathbb{E}\left(\int_0^t \sigma R(u)\sqrt{R(u)}dW(u)\right) = 0.$$

Differentiating the both sides, we have equations:

$$d\mathbb{E}R^2(t) = [(2\alpha + \sigma^2)\mathbb{E}R(t) - 2\beta\mathbb{E}R^2(t)]dt, \; \mathbb{E}R^2(0) = r_0^2.$$

Denote by $\mathbb{E}R^2(t) = y(t)$. We have

$$y'(t) = (2\alpha + \sigma^2)\mathbb{E}R(t) - 2\beta y(t), \; y(0) = r_0^2, \tag{3.6}$$

where $\mathbb{E}R(t)$ given in (3.5). By multiplying $e^{2\beta t}$ on each side of (3.6), we obtain that

$$e^{2\beta t}y'(t) + 2\beta e^{2\beta t}y(t) = e^{2\beta t}(2\alpha + \sigma^2)\mathbb{E}(R(t)).$$

Thus,

$$[e^{2\beta t}y(t)]' = e^{2\beta t}(2\alpha + \sigma^2)\mathbb{E}(R(t)).$$

Integrating the both sides, we have

$$e^{\beta t}y(t) = \int_0^t e^{2\beta s}(2\alpha + \sigma^2)\mathbb{E}(R(s))ds + C,$$

for some constant C. Thus, we have that

$$y(t) = \mathbb{E}R^2(t) = e^{-2\beta t}r_0^2 + \frac{2\alpha + \sigma^2}{\beta}\left(r_0 - \frac{\alpha}{\beta}\right)(e^{-\beta t} - e^{-2\beta t}) + \frac{\alpha(2\alpha + \sigma^2)}{2\beta^2}(1 - e^{-2\beta t}).$$

Hence, we have

$$\text{var}R(t) = \mathbb{E}R^2(t) - (\mathbb{E}R(t))^2 = \frac{\sigma^2}{\beta}(e^{-\beta t} - e^{-2\beta t}) + \frac{\alpha\sigma^2}{2\beta^2}(1 - 2e^{-\beta t} + e^{-2\beta t}).$$

3.4 Solving the general one dimensional linear SDE

Consider the general one dimensional linear SDE of the following form

$$dX(t) = [a(t) + b(t)X(t)]dt + [\gamma(t) + \sigma(t)X(t)]dW(t), \ X(0) = X_0. \tag{3.7}$$

where $\{a(t)\}, \{b(t)\}, \{\gamma(t)\}$, and $\{\sigma(t)\}$ are stochastic processes (not necessarily to be constants or deterministic functions) adapted to the filtration $\mathcal{F}(t)$.

Applying the Itô formula, we will show that

$$X(t) = Y(t)Z(t),$$

where

$$Y(t) = \exp\left(\int_0^t \left[b(s) - \frac{1}{2}\sigma(s)^2\right] ds + \int_0^t \sigma(s)dW(s)\right)$$

and

$$Z(t) = X_0 + \int_0^t [a(s) - \gamma(s)\sigma(s)]\, Z(s)^{-1}ds + \int_0^t \gamma(s)Z(s)^{-1}dW(s).$$

We note that

$$\frac{dY(t)}{Y(t)} = b(t)dt + \sigma(t)dW(t).$$

Indeed, by using the Itô product formula, we obtain that

$$\begin{aligned}
dX(t) &= Z(t)dY(t) + Y(t)dZ(t) + dY(t)dZ(t) \\
&= Z(t)\,[b(t)Y(t)dt + \sigma(t)Y(t)dW(t)] \\
&\quad + Y(t)\left[[a(t) - \gamma(t)\sigma(t)]\, Y(t)^{-1}dt + \gamma(t)Y(t)^{-1}dW(t)\right] \\
&\quad + \sigma(t)Y(t)\gamma(t)Y(t)^{-1}dt \\
&= b(t)X(t)dt + \sigma(t)X(t)dW(t) \\
&\quad + [[a(t) - \gamma(t)\sigma(t)]\, dt + \gamma(t)dW(t)] + \sigma(t)\gamma(t)dt \\
&= [a(t) + b(t)X(t)]dt + [\gamma(t) + \sigma(t)X(t)]dW(t).
\end{aligned}$$

Assuming $a(t), b(t), \gamma(t)$, and $\sigma(t)$ are all deterministic functions, we can obtain $\mathbb{E}X(t))$ and $\mathrm{Var}(X(t))$ explicitly. Integration of the both sides of (3.7) yields that

$$X(t) = X_0 + \int_0^t [a(s) + b(s)X(s)]ds + \int_0^t [\gamma(s) + \sigma(s)X(s)]dW(s)$$

Taking expectation on the both sides, we obtain

$$\begin{aligned}
\mathbb{E}X(t) &= X_0 + \int_0^t [a(s)ds + b(s)\mathbb{E}X(s)]ds + \mathbb{E}\left(\int_0^t [\gamma(s) + \sigma(s)X(s)]dW(s)\right) \\
&= X_0 + \int_0^t [a(s)ds + b(s)\mathbb{E}X(s)]ds,
\end{aligned}$$

which is equivalent to an ODE initial value problem:

$$d\mathbb{E}X(t) = a(t)dt + b(t)\mathbb{E}X(t)dt,\ \mathbb{E}X(0) = X_0.$$

Using the integrating factor $\exp\left(-\int_0^t b(s)ds\right)$, we solve it as

$$\mathbb{E}X(t) = X_0 \exp\left(\int_0^t b(s)ds\right) + \int_0^t a(s) \exp\left(\int_s^t b(u)du\right) ds.$$

Because of

$$\mathrm{var}X(t) = \mathbb{E}X^2(t) - (\mathbb{E}X(t))^2,$$

we just need to calculate $\mathbb{E}X^2(t)$. For this purpose, we use the Itô formula to obtain the dynamics of $dX^2(t)$. Then, by taking expectation and using the previous result, it is straightforward to obtain $\mathbb{E}X^2(t)$ as the solution to an ODE initial value problem. We suggest it as an excellent exercise.

Here is an interesting question to consider. How can we propose a multidimensional linear SDE properly? Under what conditions, if any, such an SDE is explicitly solvable?

4 Lecture 9

4.1 Connect Brownian motion and backward heat equation

Consider the heat equation

$$u_\tau(\tau, x) = \frac{1}{2} u_{zz}(\tau, x),$$

for all $\tau \in [0, +\infty)$ and $x \in \mathcal{R}$ with the initial condition

$$u(0, x) = f(x),$$

where $f(x)$ is continuous and uniformly bounded. We have proved that it has a unique solution:

$$u(t, x) = \int_{-\infty}^{+\infty} f(y)G(x, y, t)dy \tag{4.1}$$

where $G(x, y, t)$ is defined as

$$G(x, y, t) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(x-y)^2}{2t}\right). \tag{4.2}$$

If we consider a function $v(t, x) := u(T - t, x)$, it is obvious that

$$\frac{\partial v}{\partial t} = -\frac{\partial u}{\partial t}, \ \frac{1}{2}\frac{\partial^2 v}{\partial x^2} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2}.$$

Thus, it is straightforward to have

$$\frac{\partial v}{\partial t} + \frac{1}{2}\frac{\partial^2 v}{\partial x^2} = 0, \tag{4.3}$$

with a terminal condition

$$v(T, x) = f(x), \tag{4.4}$$

On the other hand, using (4.1) , we find that $v(t, x)$ admits the following explicit solution

$$v(t, x) = u(T - t, x) = \int_{-\infty}^{+\infty} f(y) \frac{1}{\sqrt{2\pi(T-t)}} \exp\left(-\frac{(x-y)^2}{2(T-t)}\right) dy. \tag{4.5}$$

For a standard one-dimensional Brownian motion $\{B(t)\}$, it is straightforward to have

$$\mathbb{E}[f(B(T))|B(t) = x] = \frac{1}{\sqrt{2\pi(T-t)}} \int_{-\infty}^{+\infty} f(y) \exp\left(-\frac{(y-x)^2}{2(T-t)}\right) dy, \tag{4.6}$$

by using the definition of Brownian motion. Thus, from (4.5) and (4.6), we obtain that

$$v(t, x) = \mathbb{E}[f(B(T))|B(t) = x]. \tag{4.7}$$

The previous discussion is based on calculus. We can obtain that, if a function $v(t, x)$ is defined as (4.7), by using the martingale property of the Brownian motion, we can show that it solves the backward heat equation (4.3) together with the terminal condition (4.4). Using the definition of conditional expectation and the Markov property of Brownian motion, it is easy to see that

$$v(t, B(t)) = \mathbb{E}[f(B(T))|B(t)] = \mathbb{E}[f(B(T))|\mathcal{F}(t)],$$

where $\{\mathcal{F}(t)\}$ denotes the filtration generated by the Brownian motion. Obviously, $\{\mathbb{E}[f(B(T))|\mathcal{F}(t)]\}$ is a martingale; so is $\{v(t, B(t))\}$. By using the Itô formula, we have

$$\begin{aligned} dv(t, B(t)) &= v_t(t, B(t))dt + v_x(t, B(t))dB(t) + \frac{1}{2}v_{xx}(t, B(t))dt \\ &= \left[v_t(t, B(t)) + \frac{1}{2}v_{xx}(t, B(t))\right] dt + v_x(t, B(t))dB(t). \end{aligned}$$

For the Itô process

$$v(t, B(t)) = v(0, 0) + \int_0^t \left[v_s(s, B(s)) + \frac{1}{2}v_{xx}(s, B(s))\right] ds + \int_0^t v_x(s, B(s))dB(s),$$

it is a martingale if and only if its dt term is zero. (It is a general result that an Itô process is a martingale if and only if its dt term is zero.) Thus, the PDE (4.3) is satisfied and obviously the terminal condition (4.4) is satisfied

$$v(T, x) = \mathbb{E}[f(B(T))|B(T) = x] = f(x).$$

4.2 A simple example for understanding the Girsanov theorem

Assume that we have a standard normal variable X on a probability space $(\Omega, \mathcal{F}, P)$. Now, define

$$Z = \exp\left(-\theta X - \frac{1}{2}\theta^2\right).$$

and create an equivalent probability measure $\widetilde{\mathbb{P}}$ via the Radon-Nykodim derivative:

$$\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}} = Z.$$

We just need to show that $\widetilde{\mathbb{P}}$ is indeed a probability measure. For this purpose, we deduce that

$$\begin{aligned}
\widetilde{\mathbb{P}}(\Omega) &= \int_{\Omega} d\widetilde{\mathbb{P}} = \int_{\Omega} Z d\mathbb{P} = \mathbb{E}Z \\
&= \mathbb{E}\exp\left(-\theta X - \frac{1}{2}\theta^2\right) = \left[\mathbb{E}\exp\left(-\theta X\right)\right]\exp\left(-\frac{1}{2}\theta^2\right) = 1.
\end{aligned}$$

The above argument can be rigorously interpreted via the perspective of real analysis or measure theory. However, it doesn't matter if you focus on applications and have never learnt anything about those theory. We note that the expectation $\mathbb{E}Z$ can be regarded as an integration of $Z(\omega)$ with respect to ω in the whole set Ω, i.e.,

$$\mathbb{E}Z = \int_{\Omega} Z d\mathbb{P}.$$

And, just convince yourself the follow analogy between classical integration and expectation considered here

$$\int_{R} f(x)dx \sim \int_{\Omega} Z d\mathbb{P}.$$

Under the probability measure $\widetilde{\mathbb{P}}$, it is amazing that $Y = X + \theta$ is a standard normal variable. In other words, $X = Y - \theta$ is a normal variable with distribution $N(-\theta, 1)$ under $\widetilde{\mathbb{P}}$. Indeed, it is sufficient to show that the moment generating function of Y under the probability measure $\widetilde{\mathbb{P}}$ is exactly that of a standard normal variable. We deduce that

$$\begin{aligned}
\mathbb{E}^{\widetilde{\mathbb{P}}} e^{\lambda Y} &= \int_{\Omega} e^{\lambda Y} d\widetilde{\mathbb{P}} \\
&= \int_{\Omega} e^{\lambda Y} \frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}} d\mathbb{P} \\
&= \mathbb{E}^{\mathbb{P}}\left[\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}} e^{\lambda Y}\right] \\
&= \mathbb{E}^{\mathbb{P}}\left[\exp\left(-\theta X - \frac{1}{2}\theta^2 + \lambda(X+\theta)\right)\right] \\
&= \mathbb{E}^{\mathbb{P}}\left[\exp\left((\lambda-\theta)X - \frac{1}{2}\theta^2 + \lambda\theta\right)\right] \\
&= e^{\frac{1}{2}(\lambda-\theta)^2 - \frac{1}{2}\theta^2 + \lambda\theta} = e^{\frac{1}{2}\lambda^2}.
\end{aligned}$$

So, since moment generating functions determine distributions, it is enough to see that under the probability measure $\widetilde{\mathbb{P}}$, $Y = X + \theta$ is a standard normal variable.

4.3 Discounted stock price under risk neutral probability

We will prove that, under risk neutral probability $\mathbb{Q}$, the discounted stock price $e^{-rt}S(t)$ is a martingale. Applying Itô product formula, we obtain that

$$\begin{aligned} de^{-rt}S(t) &= e^{-rt}dS(t) + S(t)de^{-rt} + de^{-rt}dS(t) \\ &= e^{-rt}\left[rS(t)dt + \sigma S(t)dW^{\mathbb{Q}}(t)\right] - re^{-rt}S(t)dt \\ &= e^{-rt}\sigma S(t)dW^{\mathbb{Q}}(t) \end{aligned}$$

Thus, we have a $\mathbb{Q}-$martingale

$$e^{-rt}S(t) = S(0) + \int_0^t e^{-rt}\sigma S(t)dW^{\mathbb{Q}}(t).$$

4.4 Derivation of the Black-Scholes-Merton Formula

Under the risk-neutral measure $\mathbb{Q}$, the Black-Scholes-Merton Model follows

$$dS(t) = rS(t)dt + \sigma S(t)dW^{\mathbb{Q}}(t).$$

For lightening notations, we simply write $W^{\mathbb{Q}}(t)$ as $W(t)$ in what follows. The time t price of a call option with strike K and maturity T admits the following representation

$$c(t, S(t)) = \mathbb{E}^{\mathbb{Q}}\left[e^{-r(T-t)}(S(T) - K)^+ | \mathcal{F}(t)\right] = \mathbb{E}^{\mathbb{Q}}\left[e^{-r(T-t)}(S(T) - K)^+ | S(t)\right]$$

where the second equality follows the Markov property of $\{S(t)\}$. In other words, we need to calculate

$$c(t, s) = \mathbb{E}^{\mathbb{Q}}\left[e^{-r(T-t)}(S(T) - K)^+ | S(t) = s\right].$$

We note that

$$(S(T)-K)^{+}=(S(T)-K)1_{\{S(T)\geq K\}}=S(T)1_{\{S(T)\geq K\}}-K1_{\{S(T)\geq K\}}.$$

We will calculate $\mathbb{E}^{\mathbb{Q}}\left[S(T)1_{\{S(T)\geq K\}}|\mathcal{F}(t)\right]$ and $\mathbb{E}^{\mathbb{Q}}\left[K1_{\{S(T)\geq K\}}|\mathcal{F}(t)\right]$.

Because

$$S(t)=S(0)\exp\left(\sigma W(t)+\left(r-\frac{1}{2}\sigma^{2}\right)t\right),$$

we have

$$\begin{aligned} S(T) &\equiv S(t)\exp\left(\sigma\left(W(T)-W(t)\right)+\left(r-\frac{1}{2}\sigma^{2}\right)\tau\right) \\ &= S(t)\exp\left(\sigma\sqrt{\tau}Y+\left(r-\frac{1}{2}\sigma^{2}\right)\tau\right), \end{aligned}$$

where Y is the standard normal random variable

$$Y=\frac{W(T)-W(t)}{\sqrt{T-t}},$$

and $\tau=T-t$.

For $\mathbb{E}^{\mathbb{Q}}\left[K1_{\{S(T)\geq K\}}|S(t)=s\right]$, we have

$$\begin{aligned} & \mathbb{E}^{\mathbb{Q}}\left[K1_{\{S(T)\geq K\}}|S(t)=s\right] \\ = & K\mathbb{E}^{\mathbb{Q}}\left[1_{\{S(T)\geq K\}}|S(t)=s\right] \\ = & K\mathbb{Q}\left(S(T)\geq K|S(t)=s\right) \\ = & K\mathbb{Q}\left(s\exp\left(\sigma\sqrt{\tau}Y+\left(r-\frac{1}{2}\sigma^{2}\right)\tau\right)\geq K|S(t)=s\right) \\ = & K\mathbb{Q}\left(s\exp\left(\sigma\sqrt{\tau}Y+\left(r-\frac{1}{2}\sigma^{2}\right)\tau\right)\geq K\right), \end{aligned}$$

where we eliminate the conditioning in the last equality because of the independence between Y and

$S(t)$. Because

$$s\exp\left(\sigma\sqrt{\tau}Y+\left(r-\frac{1}{2}\sigma^2\right)\tau\right)\geq K \Longleftrightarrow Y\geq\frac{1}{\sigma\sqrt{\tau}}\left[\log\left(\frac{K}{s}\right)-\left(r-\frac{1}{2}\sigma^2\right)\tau\right],$$

we have

$$\begin{aligned}
& \mathbb{Q}\left(s\exp\left(\sigma\sqrt{\tau}Y+\left(r-\frac{1}{2}\sigma^2\right)\tau\right)\geq K\right)\\
= & \ \mathbb{Q}\left(Y\geq\frac{1}{\sigma\sqrt{\tau}}\left[\log\left(\frac{K}{s}\right)-\left(r-\frac{1}{2}\sigma^2\right)\tau\right]\right)\\
= & \ 1-N\left(\frac{1}{\sigma\sqrt{\tau}}\left[\log\left(\frac{K}{s}\right)-\left(r-\frac{1}{2}\sigma^2\right)\tau\right]\right)\\
= & \ N\left(-\frac{1}{\sigma\sqrt{\tau}}\left[\log\left(\frac{K}{s}\right)-\left(r-\frac{1}{2}\sigma^2\right)\tau\right]\right)\\
= & \ N\left(\frac{1}{\sigma\sqrt{\tau}}\left[\log\left(\frac{s}{K}\right)+\left(r-\frac{1}{2}\sigma^2\right)\tau\right]\right).
\end{aligned}$$

Thus, it is easy to find

$$\mathbb{E}^{\mathbb{Q}}\left[K1_{\{S(T)\geq K\}}|S(t)=s\right]=KN\left(d_-\left(\tau,s\right)\right),$$

where

$$d_-(\tau,s)=\frac{1}{\sigma\sqrt{\tau}}\left[\log\left(\frac{s}{K}\right)+\left(r-\frac{1}{2}\sigma^2\right)\tau\right].$$

For $\mathbb{E}^{\mathbb{Q}}\left[S(T)1_{\{S(T)\geq K\}}|S(t)=s\right]$, we note that

$$\begin{aligned}
& \mathbb{E}^{\mathbb{Q}}\left[S(T)1_{\{S(T)\geq K\}}|S(t)=s\right] \\
= & \mathbb{E}^{\mathbb{Q}}\left[s\exp\left(\sigma\sqrt{\tau}Y+\left(r-\frac{1}{2}\sigma^2\right)\tau\right)1_{\{s\exp\left(\sigma\sqrt{\tau}Y+\left(r-\frac{1}{2}\sigma^2\right)\tau\right)\geq K\}}|S(t)=s\right] \\
= & \int_{\frac{1}{\sigma\sqrt{\tau}}\left[\log\left(\frac{K}{s}\right)-\left(r-\frac{1}{2}\sigma^2\right)\tau\right]}^{\infty} s\exp\left(\sigma\sqrt{\tau}y+\left(r-\frac{1}{2}\sigma^2\right)\tau\right)\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y^2}dy \\
= & s\exp(r\tau)\int_{\frac{1}{\sigma\sqrt{\tau}}\left[\log\left(\frac{K}{s}\right)-\left(r-\frac{1}{2}\sigma^2\right)\tau\right]}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}y^2+\sigma\sqrt{\tau}y-\frac{1}{2}\sigma^2\tau\right)dy \\
= & s\exp(r\tau)\int_{\frac{1}{\sigma\sqrt{\tau}}\left[\log\left(\frac{K}{s}\right)-\left(r-\frac{1}{2}\sigma^2\right)\tau\right]}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(y-\sigma\sqrt{\tau}\right)^2\right)dy \\
= & s\exp(r\tau)\int_{\frac{1}{\sigma\sqrt{\tau}}\left[\log\left(\frac{K}{s}\right)-\left(r-\frac{1}{2}\sigma^2\right)\tau\right]-\sigma\sqrt{\tau}}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}z^2\right)dz \\
= & s\exp(r\tau)\int_{\frac{1}{\sigma\sqrt{\tau}}\left[\log\left(\frac{K}{s}\right)-\left(r+\frac{1}{2}\sigma^2\right)\tau\right]}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}z^2\right)dz \\
= & s\exp(r\tau)\left[1-N\left(\frac{1}{\sigma\sqrt{\tau}}\left[\log\left(\frac{K}{s}\right)-\left(r+\frac{1}{2}\sigma^2\right)\tau\right]\right)\right] \\
= & s\exp(r\tau)N\left(-\frac{1}{\sigma\sqrt{\tau}}\left[\log\left(\frac{K}{s}\right)-\left(r+\frac{1}{2}\sigma^2\right)\tau\right]\right) \\
= & s\exp(r\tau)N\left(d_+(\tau,s)\right),
\end{aligned}$$

where

$$d_+(\tau,s)=\frac{1}{\sigma\sqrt{\tau}}\left[\log\left(\frac{s}{K}\right)+\left(r+\frac{1}{2}\sigma^2\right)\tau\right].$$

Therefore, we finally have

$$\begin{aligned}
c(t,s) &= \mathbb{E}^{\mathbb{Q}}\left[e^{-r(T-t)}(S(T)-K)^+|S(t)=s\right] \\
&= e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left[e^{-r(T-t)}S(T)1_{\{S(T)\geq K\}}|S(t)=s\right]-e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left[K1_{\{S(T)\geq K\}}|S(t)=s\right] \\
&= sN\left(d_+\left(\tau,s\right)\right)-e^{-r\tau}KN\left(d_-\left(\tau,s\right)\right).
\end{aligned}$$

5 Lecture 10

5.1 A long-Gamma volatility arbitrage strategy

Suppose the true dynamics for the underlying stock price is

$$dS(t) = \alpha(t)S(t)dt + \beta(t)S(t)dW(t), \tag{5.1}$$

where $\alpha(t)$ and $\beta(t)$ are two stochastic processes. Here, $\alpha(t)$ represents the return of the stock and $\beta(t)$ represents the volatility. We introduce a volatility arbitrage strategy if the implied volatility of an option with maturity T and payoff $p(S(T))$ (e.g., $(S(T) - K)^+$) is higher than the true volatility. At time zero, a trader observe the market price. By using the Black-Scholes-Merton formula, she/he calculates the implied volatility as σ_{imp} by matching the observed price at the formula. In her/his view, the price of such an option at any time between time 0 and time T, denoted by $V(t) = c(t, S(t))$, were governed by the Black-Scholes-Merton PDE, i.e.,

$$rc(t, x) = c_t(t, x) + rxc_x(t, x) + \frac{1}{2}\sigma_{imp}^2 x^2 c_{xx}(t, x), \tag{5.2}$$

with terminal condition $c(T, s) = p(s)$. We note that, in particular, $c(0, S(0))$ is the observed price. Her/his strategy includes

1. a long position of a self-financing hedging portfolio with initial value $X(0) = c(0, S(0))$ (the market price of the option) and value $X(t)$ at time t by holding $\Delta(t) = \partial c(t, S(t))/\partial x$ shares of the underlying stock $S(t)$ and investing the rest $X(t) - \Delta(t)S(t)$ in the riskless money market,
2. a short position the aforementioned option.

If the price really followed the Black-Scholes-Merton model, such a portfolio perfectly hedges to zero all the time. However, this is not the case because the real dynamics of the stock price (5.1). We will provide the following analysis. First, we note that the change of value of the aforementioned self-financing portfolio satisfies that

$$dX(t) = \Delta(t)dS(t) + r\left(X(t) - \Delta(t)S(t)\right)dt,\;\; X(0) = c(0, S(0)).$$

Second, at time T, the total profit-and-loss (P&L) of this trading strategy is given by

$$X(T) - p(S(T)) \equiv X(T) - c(T, S(T)).$$

To calculate this quantity, we denote by

$$Y(t) = X(t) - c(t, S(t)).$$

(Note that $c(t, S(t))$ is not a price! We just use it as a tool for calculation here!) Obviously, $Y(0) = 0$. Now, we use the Itô formula to study the above quatity. In particular, we note that

$$\begin{aligned}
& dY(t) \\
= \;& dX(t) - dc(t, S(t)) \\
= \;& \Delta(t)dS(t) + r\left(X(t) - \Delta(t)S(t)\right)dt - c_t(t, S(t))dt - c_x(t, S(t))dS(t) - \frac{1}{2}c_{xx}(t, S(t))d[S, S](t) \\
= \;& c_x(t, S(t))dS(t) + r\left(X(t) - \Delta(t)S(t)\right)dt - c_t(t, S(t))dt - c_x(t, S(t))dS(t) - \frac{1}{2}c_{xx}(t, S(t))\beta(t)^2S(t)^2dt \\
= \;& r\left(X(t) - c_x(t, S(t))S(t)\right)dt - c_t(t, S(t))dt - \frac{1}{2}c_{xx}(t, S(t))\beta(t)^2S(t)^2dt \\
= \;& \left[rX(t) - rc_x(t, S(t))S(t) - c_t(t, S(t)) - \frac{1}{2}c_{xx}(t, S(t))\beta(t)^2S(t)^2\right]dt \\
= \;& \left[rX(t) - rc(t, S(t)) + \frac{1}{2}\sigma_{imp}^2S(t)^2c_{xx}(t, S(t)) - \frac{1}{2}c_{xx}(t, S(t))\beta(t)^2S(t)^2\right]dt \\
= \;& rY(t)dt + \frac{1}{2}S^2(t)\frac{\partial^2 c}{\partial x^2}\left(\sigma_{imp}^2 - \beta(t)^2\right)dt, \qquad (5.3)
\end{aligned}$$

where we have used the PDE (5.2). Next, we solve $Y(T)$ from the ODE

$$dY(t) = rY(t)dt + \frac{1}{2}S^2(t)\frac{\partial^2 c}{\partial x^2}\left(\sigma_{imp}^2 - \beta(t)^2\right)dt.$$

Next, we solve for $e^{-rt}Y(t)$. Using the Itô product formula, we note that

$$\begin{aligned} d\left[e^{-rt}Y(t)\right] &= e^{-rt}dY(t) + Y(t)de^{-rt} + dY(t)de^{-rt} \\ &= e^{-rt}dY(t) + Y(t)de^{-rt} \\ &= e^{-rt}dY(t) - re^{-rt}Y(t)dt. \end{aligned}$$

Thus, we have

$$\begin{aligned} d\left[e^{-rt}Y(t)\right] &= e^{-rt}\left[rY(t)dt + \frac{1}{2}S^2(t)\frac{\partial^2 c}{\partial x^2}\left(\sigma_{imp}^2 - \beta(t)^2\right)dt\right] - re^{-rt}Y(t)dt \\ &= \frac{1}{2}e^{-rt}S^2(t)\frac{\partial^2 c}{\partial x^2}\left(\sigma_{imp}^2 - \beta(t)^2\right)dt \end{aligned}$$

Integrating the both sides, we obtain that

$$e^{-rT}Y(T) - Y(0) = \frac{1}{2}\int_0^T e^{-rt}S^2(t)\frac{\partial^2 c}{\partial x^2}\left(\sigma_{imp}^2 - \beta(t)^2\right)dt.$$

Hence, we have

$$Y(T) = \frac{1}{2}\int_0^T e^{-r(T-t)}S^2(t)\frac{\partial^2 c}{\partial x^2}\left(\sigma_{imp}^2 - \beta(t)^2\right)dt.$$

If the implied volatility σ_{imp} is higher than the spot volatility $\beta(t)$, the trader makes a positive profit due to the positive Gamma $\frac{\partial^2 c}{\partial x^2}$.