Advanced Financial Models

Chen Wei

November 1, 2023

Preface

This is the note of the course *Advanced financial models* in 2023 spring, which covers a large fraction of the content in the book *Stochastic Calculus in Finance II* by Steven Shreve. Most of the content is based on the presentation of the instructor and written down in my understanding, so there are unavoidable mistakes and typos in the note. The note presumes basic knowledge in measure-based probability theory and discrete-time martingales, so the related contents in the lectures and the book are skipped or just stated without proof. Also some technical proofs in the book are also removed. Topics of the note include: basic probability and stochastic calculus, risk-neutral pricing theory, pricing of exotic options under Black-Scholes model, brief introductions to term-structure models and jump models.

Contents

1 Review of probability and stochastic calculus

1.1 Brownian motions

Definition 1.1. A stochastic process $\{W_t\}$ defined on $t \in (0, \infty)$ is called a Brownian motion if it satisfies the following:
(1) $W_0 = 0$.
(2) W_t as a function of t is continuous a.s.
(3) W_t has independent increments: for any $0 \le t_0 \le t_1 \ldots \le t_n$, $W_{t_n} - W_{t_n-1}, \ldots, W_{t_1} - W_{t_0}, W_{t_0}$ are independent.
(4) For any $0 \le s < t$, $W_t - W_s \sim N(0, t-s)$.

The Brownian motion can be constructed as the limit a scaled symmetric random walk: let $M_k = \sum_{i=1}^k X_i$ be a simple random walk with $\mathbb{P}(X_i = -1) = \mathbb{P}(X_i = 1) = \frac{1}{2}$. Then we scale it as $W_t^{(n)} = \frac{1}{\sqrt{n}} M_{[nt]}$, where $[x]$ here is the largest integer that do not exceed x. Brownian motion can be constructed as $W_t = \lim_{n\to\infty} W_t^{(n)}$.

Theorem 1.2. Some path properties of Brownian motion is as follows:
(a) It is Holder continuous almost surely: for any $\gamma < \frac{1}{2}$ and $0 \le s \le t$, $|W_t - W_s| \le C|t-s|^\gamma$ almost surely.
(b) It oscillates near 0: the zero set $R(\omega) = \{t \ge 0 : W_t(\omega) = 0\}$ is dense almost surely.
(c) $\lim_{t\to\infty} \frac{W_t}{t} = 0$ a.s. and $\lim_{t\to 0} \frac{|W_t|}{\sqrt{2t \ln|\ln t|}} = 1$ a.s.

We can also define the p-th order variation similar to the total variation function in real analysis:

Definition 1.3. Consider a finite interval $[0, t]$ and its partition $\Pi = \{t_0, t_1, \ldots, t_n\}$. Let $\{X_t\}_{t\ge 0}$ be a stochastic process. The p-th variation of X over Π is defined to be

$$V_t^{(p)}(\Pi, X) = \sum_{k=1}^n \left| X_{t_k} - X_{t_{k-1}} \right|^p$$

If $V_t^{(p)}(\Pi, X)$ converges as $\|\Pi\| \to 0$, we call the limit as the p-th variation of X. In particular, when $p = 2$, we call it as the quadratic variation and denote it by $[X, X]_t$; when $p = 1$, it is called the total variation or first order variation.

Theorem 1.4. Let W_t be a Brownian motion, then $[W, W]_t = t$ a.s., and the first order variation is ∞ a.s..

Remark. The fact that Brownian motions have infinite total variation implies that it is not differentiable at any $t \ge 0$.

Definition 1.5. The covariation of two stochastic process is defined as (if the limit exists)

$$[f, g]_t = \lim_{\|\Pi\|\to 0} \sum_{i=0}^{n-1} (f(t_i) - f(t_{i-1}))(g(t_i) - g(t_{i-1}))$$

Here $\Pi = \{t_0, t_1, \cdots, t_n\}$ is a partition of $[0, t]$.

Theorem 1.6. Let $\{W_t\}$ be a Brownian motion. Then $[W, t]_t = 0$ a.s. and $[t, t]_t = 0$. For any two independent Brownian motions $\{W_t^{(1)}\}$ and $\{W_t^{(2)}\}$, $[W^{(1)}, W^{(2)}] = 0$.
We can also write these propositions in differential forms: $dW_t dW_t = t$, $dW_t dt = 0$, $dtdt = 0$, $dW^{(1)} dW^{(2)} = 0$.

Definition 1.7. For a given Brownian motions on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we can define a filtration $\{\mathcal{F}_t\}$ with respect to this Brownian motion which is an increasing sequence of σ-algebras satisfying the following:
(a) For any $t \ge 0$, W_t is measurable w.r.t. $\mathcal{F}_t$. (b) For any $0 \le s \le t$, $W_t - W_s$ is independent of $\mathcal{F}_s$.

Remark. Given the Brownian motion B_t, the smallest filtration would be given by $\mathcal{F}_t = \sigma(B_s : 0 \le s \le t)$, but it is not the only filtration for this brownian motion.

After defining the filtration, we can consider the distribution properties of Brownian motions:

Theorem 1.8. Suppose $\{W_t\}$ is a Brownian motion with filtration $\{F_t\}$, then it satisfies the following:
(a) It is self similar: $X_t = \frac{1}{\sqrt{a}} W_{at}$ is also a Brownian motion.
(b) $Cov(W_s, W_t) = s \wedge t$.
(c) It is a martingale: for any $0 \le s \le t$, $\mathbb{E}[W_t | \mathcal{F}_s] = W_s$.

(d) It is a Markov process: for any $\mathcal{F}_s$-measurable function h, there exists another $\mathcal{F}_s$-measurable function g s.t. $\mathbb{E}[h(W_t)|\mathcal{F}_s] = g(s, W_s)$.

1.2 More on Brownian motions

We first calculate the distribution of the Brownian motion with m sample points by specifying its characteristic function:

Theorem 1.9. Let $\{W_t\}$ be a standard Brownian motion. The moment-generating function of the random vector $(W(t_1), W(t_2), \ldots, W(t_m))$ is:

$$\begin{aligned}
&\varphi(u_1, u_2, \ldots, u_m) \\
&= \mathbb{E}\exp\{u_m W(t_m) + u_{m-1} W(t_{m-1}) + \cdots + u_1 W(t_1)\} \\
&= \exp\left\{\frac{1}{2}(u_1 + u_2 + \cdots + u_m)^2 t_1 + \frac{1}{2}(u_2 + u_3 + \cdots + u_m)^2 (t_2 - t_1) + \right. \\
&\quad \left. \cdots + \frac{1}{2}(u_{m-1} + u_m)^2 (t_{m-1} - t_{m-2}) + \frac{1}{2} u_m^2 (t_m - t_{m-1})\right\}.
\end{aligned}$$

Proof. By definition,

$$\begin{aligned}
&\varphi(u_1, u_2, \ldots, u_m) \\
&= \mathbb{E}\exp\{u_m W(t_m) + u_{m-1} W(t_{m-1}) + \cdots + u_1 W(t_1)\} \\
&= \mathbb{E}\exp\{u_m (W(t_m) - W(t_{m-1})) + (u_{m-1} + u_m)(W(t_{m-1}) - W(t_{m-2})) + \\
&\cdots + (u_1 + u_2 + \cdots + u_m) W(t_1)\} \\
&= \mathbb{E}\exp\{u_m (W(t_m) - W(t_{m-1}))\} \\
&\cdot \mathbb{E}\exp\{(u_{m-1} + u_m)(W(t_{m-1}) - W(t_{m-2}))\} \\
&\quad \cdot \mathbb{E}\exp\{(u_1 + u_2 + \cdots + u_m) W(t_1)\} \\
&= \exp\left\{\frac{1}{2} u_m^2 (t_m - t_{m-1})\right\} \cdot \exp\left\{\frac{1}{2}(u_{m-1} + u_m)^2 (t_{m-1} - t_{m-2})\right\} \\
&\cdots \exp\left\{\frac{1}{2}(u_1 + u_2 + \cdots + u_m)^2 t_1\right\} \\
&= \exp\left\{\frac{1}{2}(u_1 + u_2 + \cdots + u_m)^2 t_1 + \frac{1}{2}(u_2 + u_3 + \cdots + u_m)^2 (t_2 - t_1) + \right. \\
&\quad \left. \cdots + \frac{1}{2}(u_{m-1} + u_m)^2 (t_{m-1} - t_{m-2}) + \frac{1}{2} u_m^2 (t_m - t_{m-1})\right\}.
\end{aligned}$$

□

We would like also consider the first passage times. Similar to the random walk case, we consider its 'generating function':

Theorem 1.10. (Exponential martingale). Let $W(t), t \geq 0$, be a Brownian motion with a filtration $\mathcal{F}(t), t \geq 0$, and let σ be a constant. The process

$$Z(t) = \exp\left\{\sigma W(t) - \frac{1}{2}\sigma^2 t\right\}$$

is a martingale.

Proof. For $0 \leq s \leq t$, we have

$$\begin{aligned}
&\mathbb{E}[Z(t) \mid \mathcal{F}(s)] \\
&= \mathbb{E}\left[\exp\left\{\sigma W(t) - \frac{1}{2}\sigma^2 t\right\} \mid \mathcal{F}(s)\right] \\
&= \mathbb{E}\left[\exp\{\sigma(W(t) - W(s))\} \cdot \exp\left\{\sigma W(s) - \frac{1}{2}\sigma^2 t\right\} \mid \mathcal{F}(s)\right] \\
&= \exp\left\{\sigma W(s) - \frac{1}{2}\sigma^2 t\right\} \cdot \mathbb{E}[\exp\{\sigma(W(t) - W(s))\} \mid \mathcal{F}(s)] \\
&= \exp\left\{\sigma W(s) - \frac{1}{2}\sigma^2 t\right\} \cdot \mathbb{E}[\exp\{\sigma(W(t) - W(s))\}]
\end{aligned}$$

Because $W(t) - W(s)$ is normally distributed with mean zero and variance $t - s$, this expected value is $\exp\left\{\frac{1}{2}\sigma^2(t-s)\right\}$.

Substituting this into the formula above, we obtain the martingale property

$$\mathbb{E}[Z(t) \mid \mathcal{F}(s)] = \exp\left\{\sigma W(s) - \frac{1}{2}\sigma^2 s\right\} = Z(s)$$

$\square$

We can now specify the distribution of the first passage time:

Definition 1.11. Let m be a real number, and define the first passage time to level m

$$\tau_m = \min\{t \geq 0; W(t) = m\}.$$

This is the first time the Brownian motion W reaches the level m. If the Brownian motion never reaches the level m, we set $\tau_m = \infty$.

Theorem 1.12. For $m \in \mathbb{R}$, the first passage time of Brownian motion to level m is finite almost surely, and the Laplace transform of its distribution is given by

$$\mathbb{E}e^{-\alpha\tau_m} = e^{-|m|\sqrt{2\alpha}} \text{ for all } \alpha > 0.$$

Proof. By optional sampling theorem, as the martingale $Z_{t\wedge r_m}$ defined above is bounded, we have

$$1 = Z(0) = \mathbb{E}Z(t \wedge \tau_m) = \mathbb{E}\left[\exp\left\{\sigma W(t \wedge \tau_m) - \frac{1}{2}\sigma^2 (t \wedge \tau_m)\right\}\right]$$

For the next step, we assume that $\sigma > 0$ and $m > 0$. In this case, the Brownian motion is always at or below level m for $t \leq \tau_m$ and so

$$0 \leq \exp\{\sigma W(t \wedge \tau_m)\} \leq e^{\sigma m}.$$

If $\tau_m < \infty$, the term $\exp\left\{-\frac{1}{2}\sigma^2 (t \wedge \tau_m)\right\}$ is equal to $\exp\left\{-\frac{1}{2}\sigma^2 \tau_m\right\}$ for large enough t. On the other hand, if $\tau_m = \infty$, then the term $\exp\left\{-\frac{1}{2}\sigma^2 (t \wedge \tau_m)\right\}$ is equal to $\exp\left\{-\frac{1}{2}\sigma^2 t\right\}$, and as $t \to \infty$, this converges to zero. We capture these two cases by writing

$$\lim_{t\to\infty} \exp\left\{-\frac{1}{2}\sigma^2 (t \wedge \tau_m)\right\} = \mathbb{I}_{\{\tau_m<\infty\}} \exp\left\{-\frac{1}{2}\sigma^2 \tau_m\right\}$$

We can now take the limit above to obtain

$$1 = \mathbb{E}\left[\mathbb{I}_{\{\tau_m<\infty\}} \exp\left\{\sigma m - \frac{1}{2}\sigma^2 \tau_m\right\}\right]$$

which is just

$$\mathbb{E}\left[\mathbb{I}_{\{\tau_m<\infty\}} \exp\left\{-\frac{1}{2}\sigma^2 \tau_m\right\}\right] = e^{-\sigma m}$$

This holds when m and σ are positive. We take the limit on both sides as $\sigma \downarrow 0$. By dominated convergence theorem,

$$\mathbb{E}\left[\mathbb{I}_{\{\tau_m<\infty\}}\right] = 1$$

or, equivalently,

$$\mathbb{P}\{\tau_m < \infty\} = 1$$

Because τ_m is finite with probability one (we say τ_m is finite almost surely), we may drop the indicator of this event to obtain

$$\mathbb{E}\left[\exp\left\{-\frac{1}{2}\sigma^2 \tau_m\right\}\right] = e^{-\sigma m}$$

By taking $\alpha = 2\sigma^2$ and notice that the Brownian motion is symmetric w.r.t. the origin, we have the desired result. $\square$

Remark. Differentiation of the formula above with respect to α results in (which is justified by the dominated convergence theorem)

$$\mathbb{E}\left[\tau_m e^{-\alpha\tau_m}\right] = \frac{|m|}{\sqrt{2\alpha}} e^{-|m|\sqrt{2\alpha}} \text{ for all } \alpha > 0.$$

Letting $\alpha \downarrow 0$, we obtain $\mathbb{E}\tau_m = \infty$ so long as $m \neq 0$.

The last topic to address is the reflection principle:

Theorem 1.13. For any standard BM $\{W_t\}$,

$$\mathbb{P}\{\tau_m \leq t, W(t) \leq w\} = \mathbb{P}\{W(t) \geq 2m - w\}, \quad w \leq m, m > 0.$$

Proof. We fix a positive level m and a positive time t. We wish to "count" the Brownian motion paths that reach level m at or before time t (i.e., those paths for which the first passage time τ_m to level m is less than or equal to t). There are two types of such paths: those that reach level m prior to t but at time t are at some level w below m, and those that exceed level m at time t. There are also Brownian motion paths that are exactly at level m at time t, but the probability of this for Brownian motion is zero. We may thus ignore this possibility.
Then for each Brownian motion path that reaches level m prior to time t but is at a level w below m at time t, there is a "reflected path" that is at level $2m - w$ at time t. This reflected path is constructed by switching the up and down moves of the Brownian motion from time τ_m onward. Of course, the probability that a Brownian motion path ends at exactly w or at exactly $2m - w$ is zero. In order to have nonzero probabilities, we consider the paths that reach level m prior to time t and are at or below level w at time t, and we consider their reflections, which are at or above $2m - w$ at time t. This leads to the key reflection equality

$$\mathbb{P}\{\tau_m \leq t, W(t) \leq w\} = \mathbb{P}\{W(t) \geq 2m - w\}, \quad w \leq m, m > 0.$$

□

The direct consequence of the reflection principle is that we can determine the distribution of r_m:

Theorem 1.14. For all $m \neq 0$, the random variable τ_m has cumulative distribution function

$$\mathbb{P}\{\tau_m \leq t\} = \frac{2}{\sqrt{2\pi}} \int_{\frac{|m|}{\sqrt{t}}}^{\infty} e^{-\frac{y^2}{2}} dy, \quad t \geq 0$$

and density

$$f_{\tau_m}(t) = \frac{d}{dt}\mathbb{P}\{\tau_m \leq t\} = \frac{|m|}{t\sqrt{2\pi t}} e^{-\frac{m^2}{2t}}, \quad t \geq 0.$$

Proof. We first consider the case $m > 0$. We substitute $w = m$ into the reflection formula (3.7.1) to obtain

$$\mathbb{P}\{\tau_m \leq t, W(t) \leq m\} = \mathbb{P}\{W(t) \geq m\}.$$

On the other hand, if $W(t) \geq m$, then we are guaranteed that $\tau_m \leq t$. In other words,

$$\mathbb{P}\{\tau_m \leq t, W(t) \geq m\} = \mathbb{P}\{W(t) \geq m\}.$$

Adding these two equations, we obtain the cumulative distribution function for τ_m :

$$\begin{aligned} \mathbb{P}\{\tau_m \leq t\} &= \mathbb{P}\{\tau_m \leq t, W(t) \leq m\} + \mathbb{P}\{\tau_m \leq t, W(t) \geq m\} \\ &= 2\mathbb{P}\{W(t) \geq m\} = \frac{2}{\sqrt{2\pi t}} \int_m^{\infty} e^{-\frac{x^2}{2t}} dx. \end{aligned}$$

We make the change of variable $y = \frac{x}{\sqrt{t}}$ in the integral, and this leads to the cdf when m is positive. If m is negative, by symmetry we have the same result. Finally, the formula for pdf is obtained by differentiating cdf with respect to t. □

The reflection principle also provide a formula to calculate the extreme of the Brownian motion:

Theorem 1.15. We define the maximum to date for Brownian motion to be

$$M(t) = \max_{0 \leq s \leq t} W(s)$$

For $m > 0$, $M(t) \geq m$ if and only if $\tau_m \leq t$. This observation permits us to rewrite the reflection equality as

$$\mathbb{P}\{M(t) \geq m, W(t) \leq w\} = \mathbb{P}\{W(t) \geq 2m - w\}, \quad w \leq m, m > 0.$$

From this, we can obtain the joint distribution of $W(t)$ and $M(t)$ as

$$f_{M(t),W(t)}(m, w) = \frac{2(2m - w)}{t\sqrt{2\pi t}} e^{-\frac{(2m-w)^2}{2t}}, \quad w \leq m, m > 0.$$

Proof. Because

$$\mathbb{P}\{M(t) \geq m, W(t) \leq w\} = \int_m^{\infty} \int_{-\infty}^{w} f_{M(t),W(t)}(x, y) dy dx$$

and

$$\mathbb{P}\{W(t) \geq 2m - w\} = \frac{1}{\sqrt{2\pi t}} \int_{2m-w}^{\infty} e^{-\frac{z^2}{2t}} dz$$

we have from the reflection principle that

$$\int_m^\infty \int_{-\infty}^w f_{M(t),W(t)}(x,y)dydx = \frac{1}{\sqrt{2\pi t}} \int_{2m-w}^\infty e^{-\frac{x^2}{2t}} dz$$

We differentiate first with respect to m to obtain

$$-\int_{-\infty}^w f_{M(t),W(t)}(m,y)dy = -\frac{2}{\sqrt{2\pi t}} e^{-\frac{(2m-w)^2}{2t}}.$$

We next differentiate with respect to w to see that

$$f_{M(t),W(t)}(m,w) = \frac{2(2m-w)}{t\sqrt{2\pi t}} e^{-\frac{(2m-w)^2}{2t}}.$$

□

1.3 Ito integral

In this section we would consider Ito integral for square-integrable stochastic processes. To simplify our notations, we would regard these processes as elements of an L^2 space with norm $\|f\| = \mathbb{E}[\int_0^t f(s)^2 ds]$.

Just like from simple functions we develop the Lebesgue integral theory, we develop the Ito integral theory from simple processes.

Definition 1.16. A simple processes $\{\phi_t\}_{t\geq 0}$ adapted w.r.t. the filtration $\mathcal{F}_t$ take the following form:

$$\phi(s,\omega) = A_0(\omega)\mathbf{1}_0(s) + \sum_{k=0}^{n-1} A_k(\omega)\mathbf{1}_{(t_k,t_{k+1}]}(s)$$

where $0 = t_0 < t_1 < \cdots < t_n = T, A_k \in \mathcal{F}_{t_k}$

Now we define the Ito integral for simple processes"

Definition 1.17. Consider adapted simple process $\{\phi_t\}_{t\geq 0}$ of the form:

$$\phi(s,\omega) = A_0(\omega)\mathbf{1}_{\{0\}}(s) + \sum_{i=0}^{n-1} A_i(\omega)\mathbf{1}_{(t_i,t_{i+1}]}(s)$$

where $0 = t_0 < t_1 < \cdots < t_n = T, A_k \in \mathcal{F}_{t_k}$. Ito Integral of $\{\phi_t\}_{t\geq 0}$ is defined as

$$\begin{aligned} I_t(\phi) &= \int_0^t \phi_s dW_s \\ &= \sum_{i=0}^{k-1} A_i \left(W_{t_{i+1}} - W_{t_i}\right) + A_k \left(W_t - W_{t_k}\right), \text{ if } t \in (t_k, t_{k+1}] \end{aligned}$$

Now we can define the Ito integral for general square-integrable processes by approximation of simple functions.

Definition 1.18. Consider adapted process $\{X_t\}_{t\geq 0}$ satisfying the square-integrability condition:

$$\mathbb{E}\left[\int_0^T X_s^2 ds\right] < \infty$$

where T is a positive constant. Ito Integral of $\{X_t\}_{t\geq 0}$ is defined as

$$I_t(X) = \int_0^t X_s dW_s = \lim_{n\to\infty} I_t(\phi_n)$$

where $\phi_n, n \geq 1$ is sequences of adapted simple processes approximating $\{X_t\}$, that is, it satisfies

$$\lim_{n\to\infty} \mathbb{E}\left[\int_0^T |\phi_n(s) - X_s|^2 ds\right] = 0$$

To help justifying approximation in L^2, we have the following theorem:

Theorem 1.19. Let $\{\phi_n\}$ be a sequence of square-integrable process. Then it approximates a constant m iff

$$\mathbb{E}\left[\int_0^t \phi_n(s)ds\right] \to m, \operatorname{var}\left[\int_0^t \phi_n(s)ds\right] \to 0$$

Several properties of Ito integral are in order:

Theorem 1.20. (a) It is linear: for any two square-integrable process $\{X_t\}$ and $\{Y_t\}$ and real numbers c_1, c_2, we have $\int_0^t (c_1 X_s + c_2 Y_s)ds = c_1 \int_0^t X_s ds + c_2 \int_0^t Y_s ds$.
(b) It is a martingale as a stochastic process of time t.

Remark. It is trivially true for simple processes and hence for general processes.
(c) Ito isometry: $\mathbb{E}\left[\left(\int_0^t \theta_s d\omega_s\right)^2\right] = \mathbb{E}\left[\int_0^t \theta_s^2 ds\right]$

Remark. Informally,

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_0^t \theta_s dW_s\right)^2\right] \\
=&\mathbb{E}\left[\int_0^t \theta_s dW_s \int_0^t \theta_u dW_u\right] \\
=&\mathbb{E}\left[\int_0^t \int_0^t \theta_s \theta_u dW_s dW_u\right] \\
=&\mathbb{E}\left[\mathbb{E}\left[\int_0^t \int_0^t \theta_s \theta_u dW_s dW_u \mid \mathcal{F}_{s\wedge u}\right]\right] \\
=&\mathbb{E}\left[\int_0^t \int_0^t \mathbb{E}\left[\theta_s \theta_u dW_s dW_u \mid F_{s\wedge u}\right]\right] \\
=&\mathbb{E}\left[\int_0^t \theta_s^2 dW_s dW_s\right] \\
=&\mathbb{E}\left[\int_0^t \theta_s^2 ds\right].
\end{aligned}$$

(d) It has quadratic variation $[I, I]_t = \int_0^t \theta_s^2 ds$ if $I_t = \int_0^t \theta_s dW_s$. In differential form, $dI_t dI_t = \theta_s^2 ds$

Definition 1.21. A stochastic process $\{X_t\}_{t\geqslant 0}$ is called an Ito process if it can be written as $X_t = X_0 + \int_0^t r_s ds + \int_0^t \theta_s dW_s$, where r_s, θ_s are adapted w.r.t. the filtration $\{\mathcal{F}_t\}_{t\geqslant 0}$ of the Brownian motion $\{W_t\}_{t\geqslant 0}$. In differential form, $dX_t = r_t dt + \theta_t dW_t$. The quadratic variation of the Ito process defined above is given by $[X, X]_t = \int_0^t \theta_s^2 ds$. In differential form, $dX_t dX_t = \theta_t^2 dt$

We can now define Ito integrals w.r.t. Ito processes:

Definition 1.22. let $\{X_t\}_{t\geqslant 0}$ be as defined above, if $\{Y_t\}_{t\geqslant 0}$ is an adapted process w.r.t. the filtration $\{\mathcal{F}_t\}$, then the integral is defined as $\int_0^t Y_s dX_s = \int_0^t Y_s r_s ds + \int_0^t Y_s \theta_s dW_s$.

The most important result in stochastic calculus is the following Ito's formula:

Theorem 1.23. Let $\{X_t\} t \geqslant 0$ be an Ito process and $f(t, x) \in C^{1,2}((\mathbb{R} \times \mathbb{R})$ (f_t and f_{xx} exists and continuous), then we have the following:

$$f\left(t, X_t\right) = f\left(0, X_0\right) + \int_0^t f_t\left(s, X_s\right) ds + \int_0^t f_x\left(s, X_s\right) dX_s + \frac{1}{2} \int_0^t f_{xx}\left(s, X_s\right) d[X, X]_s$$

In differential form, $df\left(t, X_t\right) = f_t\left(t, X_t\right) dt + f_x\left(t, X_t\right)\, dX_t + \frac{1}{2} f_{xx}\left(t, X_t\right) dX_t dX_t$

As a simple application of Ito's formula, we have 'integration by parts':

Theorem 1.24. ('Integration by parts') let $f \in C(\mathbb{R})$ of bounded variation. Then

$$\int_0^t f(s) dw_s = f(t) dw_t - \int_0^t w_s df(s)$$

Example. Let $Y_t = \exp\left(-\frac{1}{2}\int_0^t \theta_s^2 ds - \int_0^t \theta_s dW_s\right)$. Then if we set $X_t = -\frac{1}{2}\int_0^t \theta_s^2 ds - \int_0^t \theta_s dW_s, Y_t = e^{X_t}$. By Ito's formula,

$$\begin{aligned} dY_t &= e^{X_t} dX_t + \frac{1}{2} e^{X_t} dX_t dX_t \\ &= e^{X_t}\left(-\frac{1}{2}\theta_t^2 dt - \theta_t dW_t\right) + \frac{1}{2} e^{X_t}\theta_t^2 dt \\ &= -e^{X_t}\theta_t dW_t = -Y_t\theta_t dW_t \\ \Rightarrow Y_t &= Y_0 - \int_0^t Y_s\theta_s dW_s = 1 - \int_0^t Y_s\theta_s dW_s \end{aligned}$$

As a result, $\{Y_t\}_{t\geqslant 0}$ is the solution to the SDE $dY_t = -Y_t\theta_t dW_t$ and hence an Ito process.

Theorem 1.25. (Levy's characterization of Brownian motions) Let $\{\omega_t\}_{t\geqslant 0}$ be a stochastic process satisfying the following:
(1) $W_0 = 0$.
(2) The path $\{W_t\}_{t\geqslant 0}$ is continuous a.s.
(3) $\{W_t\}_{t\geqslant 0}$ is a martingale.
(4) $[W, W]_t = t$.
Then $\{W_t\}_{t\geqslant 0}$ is a Brownian motion.

We can generalize things to high dimensional spaces:

Definition 1.26. (Multi-dimensional Brownian Motion) A d-dimensional stochastic process $\vec{W}_t = \left(W_t^{(1)}, \ldots, W_t^{(d)}\right)$ is said to be a d-dimensional BM if it satisfies the following:
(1) $\vec{W}_0 = 0$.
(2) The path of $\left\{\vec{W}_t\right\}_{t\geqslant 0}$ is continuous a.s.
(3) $\forall 0 \leq s < t, \vec{W}_t - \vec{W}_s \sim N\left(0, (t-s)I_d\right)$.
(4) The increments are independent.

Remark. It can be constructed by putting d independent Brownian motions in a vector.

Theorem 1.27. (Levy's characterization of d-dimensional Brownian notions) Let $\{W_t\}_{t\leq 0}$ be a d-dimensional stochastic process satisfying the following:
(1) $\vec{W}_0 = \overrightarrow{0}$.
(2) The path $\left\{\vec{W}_t\right\}_{t\geqslant 0}$ is continuous a.s.
(3) $\left\{\vec{W}_t\right\}_{t\geqslant 0}$ is a martingale.
(4) $\forall i, j \in \{1, 2, \cdots, d\}$, $\left[W^{(i)}, W^{(j)}\right]_t = t\delta_{ij}$.
Then $\{W_t\}_{t\geqslant 0}$ is a Brownian motion.

Definition 1.28. (Muti-dimensional Ito process) Let $\vec{X}_t = \left(X_t^{(1)}, \cdots, X_t^{(n)}\right)$ be a n-dimensional stochastic process. It is called a Ito process if it can be mitten as

$$\vec{X}_t = \vec{X}_0 + \int_0^t \vec{r}_s ds + \int_0^t \vec{\theta}_s d\vec{W}_s$$

where $\vec{r}_s = \left(r_s^{(1)}, \ldots, r_s^{(n)}\right)$ is an n-dimensional process adapted to $\{\mathcal{F}_t\}_{t\geq 0}$ and $\vec{\theta}_s$ is an $n \times d$ matrix with stochastic processes as its entries.

Theorem 1.29. (Multi-dimensional Ito formula) Suppose $f(t, \vec{x}) : [0, \infty) \times \mathbb{R}^n \to \mathbb{R}$ has a continuous derivative w.r.t. t and has continuous second partial derivatives w.r.t. $\vec{x}$, then for any Ito process $\{X_t\}_{t\geqslant 0}$, we have

$$f\left(t, \vec{X}_t\right) = f\left(0, \vec{X}_0\right) + \int_0^t f\left(s, \vec{X}_s\right) ds + \int_0^t \nabla_x f\left(s, \vec{X}_s\right) d\vec{X}_s + \frac{1}{2}\int_0^t \left(d\vec{X}_s\right)^\top \nabla_{xx} f\left(s, \vec{X}_s\right) d\vec{X}_s$$

where the last term should be understood as

$$\sum_{i=1}^n \sum_{j=1}^n \int_0^t \frac{\partial^2}{\partial x_i x_j} f\left(s, \vec{X}_s\right) d\left[X_s^{(i)}, X_s^{(j)}\right]$$

In differential form,

$$df\left(t,\vec{X}_t\right)=f\left(t,\vec{X}_t\right)dt+\nabla_x f\left(t,\vec{X}_t\right)^{\top}d\vec{X}_t+\frac{1}{2}\left(d\vec{X}_t\right)^{\top}\nabla_{xx}f\left(t,\vec{X}_t\right)\left(d\vec{x}_t\right)^{\top}$$

A direct consequence is the following product rule similar to that in ordinary calculus:

Theorem 1.30. (Product rule) Let $\{X_t\}$, $\{Y_t\}$ be two Ito processes. Then

$$d\left[X_tY_t\right]=\left(dX_t\right)Y_t+\left(dY_t\right)X_t+\left(dX_t\right)\left(dY_t\right)$$

Proof. Let $f(t,\vec{X})=f\left(t,X_t,Y_t\right)=X_tY_t$, then $f_t(t,\vec{X})=0,\nabla_x f(t,\vec{X})=\begin{pmatrix}Y_t\\X_t\end{pmatrix},\nabla_{xx}f(t,\vec{X})=\begin{pmatrix}0&1\\1&0\end{pmatrix}$. By Multi-dimensional Ito formula,

$$\begin{aligned}d\left[X_tY_t\right]&=f_tdt+\left(\nabla_x f\right)^{\top}\begin{pmatrix}dX_t\\dY_t\end{pmatrix}+\frac{1}{2}\left(dX_t\ dY_t\right)\left(\nabla_{xx}f\right)\begin{pmatrix}dX_t\\dY_t\end{pmatrix}\\&=Y_tdX_t+X_tdY_t+dX_tdY_t.\end{aligned}$$

□

1.4 Brownian bridge

We move on to the discussion of the Brownian bridge. This is a stochastic process that is like a Brownian motion except that with probability one it reaches a specified point at a specified positive time. We first discuss Gaussian processes in general, the class to which the Brownian bridge belongs, and we then define the Brownian bridge and present its properties.

Definition 1.31. A Gaussian process $X(t),t\geq 0$, is a stochastic process that has the property that, for arbitrary times $0<t_1<t_2<\cdots<t_n$, the random variables $X\left(t_1\right),X\left(t_2\right),\ldots X\left(t_n\right)$ are jointly normally distributed.

Example. (Ito integral of a deterministic integrand). Let $\Delta(t)$ be a nonrandom function of time, and define

$$I(t)=\int_0^t\Delta(s)dW(s),$$

where $W(t)$ is a Brownian motion. Then $I(t)$ is a Gaussian process.

Definition 1.32. Let $W(t)$ be a Brownian motion. Fix $T>0$. We define the Brownian bridge from 0 to 0 on $[0,T]$ to be the process

$$X(t)=W(t)-\frac{t}{T}W(T),0\leq t\leq T.$$

Theorem 1.33. The Brownian bridge define above is a Gaussian process.

Proof. For $0<t_1<t_2<\cdots<t_n<T$, the random variables

$$X\left(t_1\right)=W\left(t_1\right)-\frac{t_1}{T}W(T),\ldots,X\left(t_n\right)=W\left(t_n\right)-\frac{t_n}{T}W(T)$$

are jointly normal because $W\left(t_1\right),\ldots,W\left(t_n\right),W(T)$ are jointly normal Hence, the Brownian bridge from 0 to 0 is a Gaussian process. Its mean function is easily seen to be

$$m(t)=\mathbb{E}X(t)=\mathbb{E}\left[W(t)-\frac{t}{T}W(T)\right]=0.$$

For $s,t\in(0,T)$, we compute the covariance function

$$\begin{aligned}&c(s,t)\\&=\mathbb{E}\left[\left(W(s)-\frac{s}{T}W(T)\right)\left(W(t)-\frac{t}{T}W(T)\right)\right]\\&=\mathbb{E}[W(s)W(t)]-\frac{t}{T}\mathbb{E}[W(s)W(T)]-\frac{s}{T}\mathbb{E}[W(t)W(T)]+\frac{st}{T^2}\mathbb{E}W^2(T)\\&=s\wedge t-\frac{2st}{T}+\frac{st}{T}=s\wedge t-\frac{st}{T}\end{aligned}$$

□

Remark. Note that process $X(t)$ satisfies

$$X(0) = X(T) = 0.$$

Because $W(T)$ enters the definition of $X(t)$ for $0 \le t \le T$, the Brownian bridge $X(t)$ is not adapted to the filtration $\mathcal{F}(t)$ generated by $W(t)$. We shall later obtain a different process that has the same distribution as the process $X(t)$ but is adapted to this filtration.

Definition 1.34. (General Brownian bridges) Let $W(t)$ be a Brownian motion. Fix $T > 0, a \in \mathbb{R}$, and $b \in \mathbb{R}$. We define the Brownian bridge from a to b on $[0, T]$ to be the process

$$X^{a \to b}(t) = a + \frac{(b-a)t}{T} + X(t), 0 \le t \le T,$$

where $X(t) = X^{0 \to 0}$ is the Brownian bridge from 0 to 0 of the definition above.

Remark. The function $a + \frac{(b-a)t}{T}$, as a function of t, is the line from $(0, a)$ to (T, b). When we add this line to the Brownian bridge from 0 to 0 on $[0, T]$, we obtain a process that begins at a at time 0 and ends at b at time T. Adding a nonrandom function to a Gaussian process gives us another Gaussian process. The mean function is affected:

$$m^{a \to b}(t) = \mathbb{E} X^{a \to b}(t) = a + \frac{(b-a)t}{T}.$$

However, the covariance function is not affected:

$$c^{a \to b}(s, t) = \mathbb{E}\left[\left(X^{a \to b}(s) - m^{a \to b}(s)\right)\left(X^{a \to b}(t) - m^{a \to b}(t)\right)\right] = s \wedge t - \frac{st}{T}$$

We cannot write the Brownian bridge as a stochastic integral of a deterministic integrand because the variance of the Brownian bridge,

$$\mathbb{E} X^2(t) = c(t, t) = t - \frac{t^2}{T} = \frac{t(T-t)}{T},$$

increases for $0 \le t \le \frac{T}{2}$ and then decreases for $\frac{T}{2} \le t \le T$. In the example above, the variance of $I(t) = \int_0^t \Delta(u) dW(u)$ is $\int_0^t \Delta^2(u) du$, which is nondecreasing in t. However, we can obtain a process with the same distribution as the Brownian bridge from 0 to 0 as a scaled stochastic integral In particular, consider

$$Y(t) = (T-t) \int_0^t \frac{1}{T-u} dW(u), 0 \le t < T.$$

The integral

$$I(t) = \int_0^t \frac{1}{T-u} dW(u)$$

is a Gaussian process, provided $t < T$ so the integrand is defined. For $0 < t_1 < t_2 < \cdots < t_n < T$, the random variables

$$Y(t_1) = (T - t_1) I(t_1), Y(t_2) = (T - t_2) I(t_2), \ldots, Y(t_n) = (T - t_n) I(t_n)$$

are jointly normal because $I(t_1), I(t_2), \ldots, I(t_n)$ are jointly normal In particular, Y is a Gaussian process. The mean and covariance functions of I are

$$\begin{aligned} m^I(t) &= 0 \\ c^I(s, t) &= \int_0^{s \wedge t} \frac{1}{(T-u)^2} du = \frac{1}{T - s \wedge t} - \frac{1}{T} \text{ for all } s, t \in [0, T). \end{aligned}$$

This means that the mean function for Y is $m^Y(t) = 0$. To compute the covariance function for Y, we assume for the moment that $0 \le s \le t < T$ so that

$$c^I(s, t) = \frac{1}{T-s} - \frac{1}{T} = \frac{s}{T(T-s)}$$

Then

$$\begin{aligned} c^Y(s, t) &= \mathbb{E}[(T-s)(T-t) I(s) I(t)] \\ &= (T-s)(T-t) \frac{s}{T(T-s)} \\ &= \frac{(T-t)s}{T} \\ &= s - \frac{st}{T} \end{aligned}$$

If we had taken $0 \le t \le s < T$, the roles of s and t would have been reversed. In general,

$$c^Y(s, t) = s \wedge t - \frac{st}{T} \text{ for all } s, t \in [0, T)$$

This is the same covariance formula as we obtained for the Brownian bridge. Because the mean and covariance functions for a Gaussian process completely determine the distribution of the process, we conclude that the process Y has the same distribution as the Brownian bridge from 0 to 0 on $[0, T]$ We now consider the variance

$$\mathbb{E}Y^2(t) = c^Y(t,t) = \frac{t(T-t)}{T}, 0 < t < T$$

Note that, as $t \uparrow T$, this variance converges to 0 . In other words, as $t \uparrow T$, the random process $Y(t)$, which always has mean zero, has a variance that converges to zero. We did not initially define $Y(T)$, but this observation suggests that it makes sense to define $Y(T) = 0$. If we do that, then $Y(t)$ is continuous at $t = T$. We summarize this discussion with the following theorem.

Theorem 1.35. Define the process

$$Y(t) = \begin{cases} (T-t)\int_0^t \frac{1}{T-u} dW(u) & \text{for } 0 \le t < T, \\ 0 & \text{for } t = T. \end{cases}$$

Then $Y(t)$ is a continuous Gaussian process on $[0, T]$ and has mean and covariance functions

$$m^Y(t) = 0, t \in [0, T],$$
$$c^Y(s,t) = s \wedge t - \frac{st}{T} \text{ for all } s, t \in [0, T].$$

In particular, the process $Y(t)$ has the same distribution as the Brownian bridge from 0 to 0 on $[0, T]$.

Remark. We note that the process $Y(t)$ is adapted to the filtration generated by the Brownian motion $W(t)$. It is interesting to compute the stochastic differential of $Y(t)$, which is

$$\begin{aligned} dY(t) &= \int_0^t \frac{1}{T-u} dW(u) \cdot d(T-t) + (T-t) \cdot d\int_0^t \frac{1}{T-u} dW(u) \\ &= -\int_0^t \frac{1}{T-u} dW(u) \cdot dt + dW(t) \\ &= -\frac{Y(t)}{T-t} dt + dW(t). \end{aligned}$$

If $Y(t)$ is positive as t approaches T, the drift term $-\frac{Y(t)}{T-t}dt$ becomes large in absolute value and is negative. This drives $Y(t)$ toward zero. On the other hand, if $Y(t)$ is negative, the drift term becomes large and positive, and this again drives $Y(t)$ toward zero. This strongly suggests, and it is indeed true, that as $t \uparrow T$ the process $Y(t)$ converges to zero almost surely.

1.5 Change of measure by random variables

Theorem 1.36. Let Z be a random variable with expectation $\mathbb{E}[Z] = 1$ in the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then for $\mathcal{F}$, we can define a new measure by $\tilde{\mathbb{P}}(A) = \mathbb{E}[Z 1_A] = \int_A Z dP$.

Proof. It follows from measure theory. □

Definition 1.37. For any two probability measures $\mathbb{P}$ and $\tilde{\mathbb{P}}$ defined on the same probability space $(\Omega, \mathcal{F})$, if there exists a random variable Z such that $\tilde{\mathbb{P}}(A) = \mathbb{E}[Z 1_A] = \int_A Z dP$, we say that Z is the Radon-Nikodym derivative of $\tilde{P}$ with respect to P.

Theorem 1.38. Suppose the expectation of the random variable X defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ exists, and Z is the Radon-Nikodym derivative of $\tilde{\mathbb{P}}$ with respect to $\mathbb{P}$, then $\tilde{\mathbb{E}}[X] = \mathbb{E}[ZX]$.
If $Z > 0$ a.s., then $\mathbb{E}[X] = \tilde{\mathbb{E}}[\frac{X}{Z}]$.

Proof. Just apply the change of variable formula. □

Definition 1.39. Let $\mathbb{P}$ and $\tilde{\mathbb{P}}$ be two probability measures defined on the probability space $(\Omega, \mathcal{F})$, we say that $\mathbb{P}$ and $\tilde{\mathbb{P}}$ are equivalent if $\mathbb{P}(A) = 0 \Leftrightarrow \tilde{\mathbb{P}}(A) = 0$ for any event $A \in \mathcal{F}$.

Remark. If $\mathbb{P}$ and $\tilde{\mathbb{P}}$ are equivalent then they agree on the events with probability zero/positive probability/probability one.

Example. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and X be a standard normal random variable. Suppose there is a random variable Z such that $Z = e^{-\theta X - \frac{1}{2}\theta^2}$, where θ is a constant and the Radon-Nikodym derivative of a probability measure $\tilde{\mathbb{P}}$ with respect to $\mathbb{P}$ is Z. Find the distribution of X under $\tilde{\mathbb{P}}$.

Solution. We just calculate its characteristics function:

$$\tilde{\mathbb{E}}[e^{itX}] = \mathbb{E}[Ze^{itX}] = \mathbb{E}[e^{-\theta X - \frac{1}{2}\theta^2 + itX}] = e^{-\frac{1}{2}\theta^2}\mathbb{E}[e^{(it-\theta)X}] = e^{-\frac{1}{2}\theta^2}e^{\frac{1}{2}(it-\theta)^2} = e^{-\frac{t^2}{2} - it\theta}$$

Hence X follows $N(-\theta, 1)$ under $\tilde{\mathbb{P}}$.

1.6 Change of measure by stochastic processes

Definition 1.40. Let $\{W_t\}_{t\geqslant 0}$ be a BM with filtration $\{\mathcal{F}_t\}$. For any time t, if $Z_t \geqslant 0$ is a random variable w.r.t. $\mathcal{F}_t$ and $\mathbb{E}[Z_t] = 1$, we can introduce a new probability measure $\widetilde{\mathbb{P}}_t(A) = \mathbb{E}[Z_t 1_A]$ on $\mathcal{F}_t$ (which is equivalent to say that $Z_t(\omega) = \frac{d\widetilde{\mathbb{P}}_t(\omega)}{d\mathbb{P}(\omega)}$). If all these measures $\left\{\widetilde{\mathbb{P}}_t\right\}$ agrees on the σ-algebra $\mathcal{F}_T$ at some terminal time T, combining all such Z_t forms a stochastic process $\{Z_t\}_{t\geqslant 0}^T$ which is called the Radon-Nikodym derivative process of $\widetilde{\mathbb{P}}$ w.r.t. $\mathbb{P}$, where $\widetilde{\mathbb{P}} \triangleq \widetilde{\mathbb{P}}_T$.

Remark. It is sufficient that $\{Z_t\}_{t\geqslant 0}^T$ is a martingale for these measures $\widetilde{\mathbb{P}}_t$ agree:

$\forall A \in \mathcal{F}_t, 0 \le t \le T, \widetilde{\mathbb{P}}_T(A) = \mathbb{E}[Z_T 1_A] = \mathbb{E}[\mathbb{E}[Z_T 1_A \mid \mathcal{F}_t]] = \mathbb{E}[1_A \mathbb{E}[Z_T \mid \mathcal{F}_t]] = \mathbb{E}[1_A Z_t] = \widetilde{\mathbb{P}}_t(A)$. So $\widetilde{\mathbb{P}}_T$ and $\widetilde{\mathbb{P}}_t$ agree.

Lemma 1.41. Let Z be a r.v. w.r.t. $\mathcal{F}_T$, where $T > 0$ is fixed. Then if we define $Z_t = \mathbb{E}[Z \mid \mathcal{F}_t]$ $\forall 0 \le t \le T$, we can introduce a new probability measure $\widetilde{\mathbb{P}}(A) = \mathbb{E}[Z_T 1_A]$ which satisfies the following:
For any r.v. Y that is $\mathcal{F}_t$-measurable, $0 \le t \le T$, we have $\widetilde{\mathbb{E}}[Y] = \mathbb{E}[YZ_t]$.

Proof. By definition,

$$\begin{aligned}\widetilde{\mathbb{E}}[Y] &= \mathbb{E}[Z_T Y] \\ &= \mathbb{E}[\mathbb{E}[Z_T Y \mid \mathcal{F}_t]] \\ &= \mathbb{E}[YE[Z_T \mid \mathcal{F}_t]] \\ &= \mathbb{E}[YZ_t].\end{aligned}$$

□

Moreover, we have a more general result under this construction.

Theorem 1.42. For any $0 \le s \le t \le T$ and a random variable Y that is $\mathcal{F}_t$-measurable, we have

$$\widetilde{\mathbb{E}}[Y \mid \mathcal{F}_s] = \frac{1}{Z_s}\mathbb{E}[YZ_t \mid \mathcal{F}_s]$$

Proof. By definition it suffices to check that $\forall A \in \mathcal{F}_s$,

$$\begin{aligned}&\widetilde{\mathbb{E}}\left[\frac{1}{Z_s}\mathbb{E}[YZ_t \mid \mathcal{F}_s]1_A\right] = \mathbb{E}\left[Z_s \cdot \frac{1}{Z_s}\mathbb{E}[YZ_t \mid \mathcal{F}_s]1_A\right] \\ =&\mathbb{E}[\mathbb{E}[YZ_t \mid \mathcal{F}_s]1_A] = \mathbb{E}[\mathbb{E}[YZ_t 1_A \mid \mathcal{F}_s]] \\ =&\mathbb{E}[YZ_t 1_A] = \widetilde{\mathbb{E}}[Y1_A].\end{aligned}$$

□

Now it comes the most important theorem in change of measure:

Theorem 1.43. (Girsanov's Theorem) Suppose $\{W_t\}_{t\geqslant 0}$ is a Brownian Motion with filtration $\{\mathcal{F}_t\}_{t\geqslant 0}$ and $\{\theta_t\}_{t\geqslant 0}$ is an adapted process w.r.t. $\{\mathcal{F}_t\}_{t\geqslant 0}$. Let $\{Z_t\}_{t\geqslant 0}^T$ be a stochastic process satisfying

$$Z_t = \exp\left(-\int_0^t \theta_u dW_u - \frac{1}{2}\int_0^t \theta_u^2 du\right) \quad \forall 0 \le t \le T$$

If $Z_T = \frac{d\widetilde{\mathbb{P}}(\omega)}{d\mathbb{P}(\omega)}$ for a new probability measure $\widetilde{\mathbb{P}}$, then the process defined by

$$\widetilde{W}_t = W_t + \int_0^t \theta_u du$$

is a Brownian Motion w.r.t. $\tilde{P}$.

Remark. $\left\{\widetilde{W}_t\right\}$ adds a random drift on $\{W_t\}$.

Proof. Let $X_t = -\int_0^t \theta_u dW_u - \frac{1}{2}\int_0^t \theta_u^2 du$, then $Z_t = e^{X_t}$, $dX_t = -\theta_t dW_t - \frac{1}{2}\theta_t^2 dt, dX_t dX_t = \theta_t^2 dt$.
Apply Ito formula for $f(t,x) = e^x$, we have

$$\begin{aligned} dZ_t = df\left(t, X_t\right) &= f_x\left(t, X_t\right) dX_t + \frac{1}{2} f_{xx}\left(t_1 X_t\right) dX_t dX_t \\ &= e^{X_t}\left(-\theta_t dW_t - \frac{1}{2}\theta_t^2 dt\right) + \frac{1}{2} e^{X_t}\theta_t^2 dt \\ &= e^{X_t}\left(-\theta_t\right) dW_t \end{aligned}$$

$\Rightarrow Z_t = Z_0 - \int_0^t e^{X_s}\theta_s dW_s$ is a martingale, because it is an Ito integral. As a result, we have $\mathbb{E}\left[Z_T\right] = \mathbb{E}\left[Z_0\right] = 1$ and $Z_T \geqslant 0$ a.s. we can define $\widetilde{\mathbb{P}}$ properly.
Now we apply Levy's characterization to show that $\widetilde{W}_t$ is a BM. Trivially $\widetilde{W}_0 = 0$ and $\widetilde{W}_t$ is continuous a.s. Moreover, $\left[\widetilde{W}_t, \widetilde{W}_t\right] = t$ as the Riemann integral does not contribute to the quadratic variation. Now we check the martingale property: $\forall 0 \leq s \leq t \leq T, \widetilde{\mathbb{E}}\left[\widetilde{W}_t \mid \mathcal{F}_s\right] = \frac{1}{Z_s}\mathbb{E}\left[\widetilde{W}_t Z_t \mid \mathcal{F}_s\right] = \frac{1}{Z_s}\widetilde{W}_s Z_s = \widetilde{W}_s$. Hence $\left\{\widetilde{W}_t\right\}_{t\geqslant 0}^T$ is a Brownian motion. $\square$

1.7 Stochastic differential equations

Definition 1.44. A stochastic differential equation is of the form

$$\left\{\begin{array}{l} X_T = X_t + \int_t^T \beta\left(u, X_u\right) du + \int_t^T \gamma\left(u, x_u\right) dW_u \\ X_t = x \end{array}\right.$$

for some deterministic functions β and γ and time $t < T$.
In differential form,

$$dX_t = \beta\left(t, X_t\right) dt + \gamma\left(t, X_t\right) dW_t.$$

Remark. Euler's discretization of SDE: for a small step time δ, on $[t, t+\delta]$ we have

$$\begin{aligned} X_{t+\delta} &= X_t + \int_t^{t+\delta} \beta\left(u, X_u\right) du + \int_t^{t+\delta} \gamma\left(u, X_u\right) dW_u \\ &\approx X_t + \int_t^{t+\delta} \beta\left(t, X_t\right) du + \int_t^{t+\delta} \gamma\left(t, X_t\right) dW_u \\ &= X_t + \beta\left(t_1, x_t\right)\delta + \gamma\left(t, X_t\right)\left(W_{t+\delta} - W_t\right). \end{aligned}$$

Base on a realized path $\{W_t\}$, we can simulate realized paths of X.

Example. Geometric Brownian motion $\{X_t\}$ is a process satisfying $dS_t = \mu S_t dt + \sigma S_t dW_t$. This SDE can be solved analytically using Ito's formula:

$$\begin{aligned} d\left[\ln S_t\right] &= \frac{1}{S_t} dS_t - \frac{1}{2}\cdot\frac{1}{S_t^2} dS_t dS_t \\ &= \mu dt + \sigma dW_t - \frac{1}{2}\cdot\frac{1}{S_t^2}\sigma^2 S_t^2 dt \\ &= \left(\mu - \frac{1}{2}\sigma^2\right) dt + \sigma dW_t \\ \Rightarrow \ln S_t - \ln S_0 &= \left(\mu - \frac{1}{2}\sigma^2\right) t + \sigma dW_t, \\ S_t &= S_0 \exp\left(\left(\mu - \frac{1}{2}\sigma^2\right) t + \sigma dW_t\right). \end{aligned}$$

Example. An OU process $\{X_t\}$ is a process satisfying the SDE

$$dR_t = \kappa\left(\theta - R_t\right) dt + \sigma dW_t$$

This SDE can also be solved analytically:

$$d\left[e^{kt} Rt\right] =$$

Example. Linear SDEs can be solved analytically:

$$dX_t = \left(a(t) + b(t) X_t\right) dt + \left(c(t) + h(t) X_t\right) dw_t$$

is given by

$$X_t = Y_t\left(X_0 + \int_0^t \frac{a(s)-c(s)h(s)}{Y_s}ds + \int_0^t \frac{c(s)}{Y_s}dW_s\right)$$

where

$$Y_t = \exp\left(\int_0^t \left(b(s) - \frac{1}{2}h(s)^2\right)ds + \int_0^t h(s)dW_s\right)$$

Proof. By Ito's formula, $\{Y_t\}$ is an inhomongeneous geometric BM satisfying

$$dY_t = b(t)Y_t + h(t)Y_t dw_t$$

(for a proof, apply Ito's formula for $\ln Y_t$)
By setting $Z(t) = \int_0^t \frac{a(s)-c(s)h(s)}{Y_s}ds + \int_0^t \frac{c(s)}{Y_s}dW_s$, we have $X_t = Y_t Z_t, dX_t = (dY_t)Z_t + (dZ_t)Y_t + d[Y_t Z_t]$.
As $dZ_t = \frac{a(t)-c(t)h(t)}{Yt}dt + \frac{d(t)}{Yt}dW_t$, we further have

$$\begin{aligned}
Y_t dZ_t &= (a(t) - c(t)h(t))dt + d(t)dW_t \\
Z_t dY_t &= b(t)Y_t Z_t dt + h(t)Y_t Z_t dW_t \\
dY_t dZ_t &= \left[\frac{1}{Y_t}(a(t) - c(t)h(t)dt + c(t)dW_t)\right][b(t)Y_t + h(t)Y_t dW_t] = c(t)h(t)dt \\
\Rightarrow dX_t &= (dY_t)Z_t + (dZ_t)Y_t + dY_t dZ_t \\
&= h(t)X_t dt + h(t)X_t dW_t + a(t)dt + c(t)dW_t.
\end{aligned}$$

□

Theorem 1.45. The solution $\{X_t\}$ to SDEs

$$dX_t = \beta(t, X_t)dt + \gamma(t, X_t)dW_t$$

are Markov processes.

By this property, we have the famous Feymann-Kac Theorem:

Theorem 1.46. Suppose $\{X_t\}$ satisfies the SDE above and f is smooth enough (s.t. f_t. f_x, f_{xx} exist and are continuous).
Then the expected value

$$f(t,x) = \mathbb{E}\left[e^{-\int_t^T \kappa(u,X_u)du}h(X_T) \mid X_t = x\right]$$

is given by the IVP of the PDE

$$\begin{cases} \frac{\partial}{\partial t}f + \beta f_x + \frac{1}{2}\gamma^2 f_{xx} - \kappa f = 0 \\ f(T,x) = h(x) \end{cases}$$

Proof. $\forall 0 \le t \le T, e^{-\int_0^t \kappa(u,x_u)du}f(t,x_t) = \mathbb{E}\left[e^{-\int_0^T \kappa(u,x_u)du}f(X_T) \mid \mathcal{F}_t\right]$ is the conditional expectation of an $\mathcal{F}_T$-measurable r.v. and hence a martingale.
By chain mole and Ito's formula,

$$\begin{aligned}
d\left[e^{-\int_0^t \kappa(u,X_u)du}\right] &= e^{-\int_0^t \kappa(u,X_u)du}(-\kappa(t,X_t))dt \\
d[f(t,X_t)] &= f_t dt + f_x dX_t + \frac{1}{2}f_{xx}dX_t dX_t \\
&= \left[f_t + \beta f_x + \frac{1}{2}f_{xx}\gamma^2\right]dt + \gamma f_x dW_t
\end{aligned}$$

Then an application of product rule yields

$$\begin{aligned}
&d\left[e^{-\int_0^t \kappa(u,x_u)du}f(t,X_t)\right] \\
=&e^{-\int_0^t \kappa(u,X_u)du}\left(f_t + \beta f_x + \frac{1}{2}f_{xx}\gamma^2\right)dt + e^{-\int_0^t \kappa(u,X_u)du}\gamma f_x dW_t - e^{-\int_0^t \kappa(u,X_u)du}\kappa(t,X_t)dtf(t,X_t) \\
=&e^{-\int_0^t k\, uX_u)du}\left(f_t + \beta f_x + \frac{1}{2}f_{xx}\gamma^2 - \kappa f\right)dt + e^{-\int_0^t \kappa(u,X_u)du}\gamma f_x dW_t
\end{aligned}$$

As it is a martingale, the drift term containing at must be zero $\Rightarrow f_t + \beta f_x + \frac{1}{2}f_{xx}\gamma^2 - \kappa f = 0$.
The initial value is attained directly by calculating the expectation at $t = T$. □

Remark. General idea of calculating expectations:
(1) Construct a martingale
(2) Apply Ito's formula to this martingale
(3) The drift term is 0, which usually gives a PDE.

Definition 1.47. Multi-dimensional SDEs are of the form

$$d\overrightarrow{X_t} = \vec{\beta}(t, \vec{X_t})dt + \gamma\left(t, \overrightarrow{X_t}\right) d\vec{W_t}$$

where $\vec{X_t}$ is an m-dimensional process, $\vec{\beta}$ is a function from $\mathbb{R} \times \mathbb{R}^m$ to $\mathbb{R}^m$, γ is a function from $\mathbb{R} \times \mathbb{R}^m$ to $\mathbb{R}^{m \times d}$ and $\widetilde{W}_t$ is a d-dimensional BM.

Its solutions $\left\{\vec{X_t}\right\}$ is still a Markov process and hence we can generalize the Feynman-Kac Theorem into multi-dimensional cases:

Theorem 1.48. Suppose $\{X_t\}$ is a solution to the SDE mentioned above,

$$f(t, x) = \mathbb{E}\left[e^{-\int_t^T \kappa\left(u, \vec{X}_u\right) du} h\left(\vec{X}_T\right) \mid \vec{X_t} = x\right]$$

and f is smooth enough (f_t, f_x, f_{xx} exists and is continuous), then it satisfies the IVP

$$\begin{aligned} & f_t + \vec{\beta}(t, x)^\top \nabla_x f + \frac{1}{2} \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial xj} \cdot \left(\gamma(t, x)\gamma(t, x)^\top\right)_{i,j} - \kappa(t, x) f = 0 \\ & f(T, x) = h(x) \end{aligned}$$

Proof. The proof is similar as the one-dimensional case. □

2 Risk-neutral Pricing

2.1 Introduction

The main purpose of this course is to set up financial models and give a fair price to them. The financial products we consider here include the following:

- **Stocks**: refers to a unit of ownership in a company. When a company goes public, it offers shares of its ownership to the public in exchange for capital. The stockholders or shareholders own a portion of the company, which entitles them to a share of the company's profits and assets. The dividends, which is a core factor in stock pricing, depend on the curriculum of the market and may suffer from volatility.
- **Bonds**: a debt security that is issued by a company, government or other entity to raise capital. When an investor buys a bond, they are effectively lending money to the issuer. The bond issuer agrees to pay interest on the bond at regular intervals and to repay the principal amount of the bond at a specified date in the future. Bonds are generally considered to be less risky than stocks, but also typically offer lower potential returns.

  Remark. The interest rate for the bond may be stochastic, which drives the randomness in bond price. However, it is usually less volatile compared to the stock price.

  Deposit: which is to keep the money in the bank and receive fixed interest.

  Remark. When the interest rate increases, the bond price usually decreases as people would prefer more to put the money in the blank. The case is similar when the interest rate decreases.

  Commodities: are physical goods that are traded on exchanges, such as gold, oil, agricultural products, and other raw materials. Commodities are often used as inputs in the production of other goods, and their prices can be influenced by a variety of factors, such as supply and demand, global economic conditions, and weather patterns. Trading in commodities can be risky, as prices can be volatile, and commodities markets can be subject to sudden changes in supply and demand.

  Financial derivatives: Financial derivatives are financial contracts that derive their value from an underlying asset, such as a stock, bond, commodity, or currency. The financial derivatives usually gives a payoff in the future and has a terminal time, called the maturity.

The main problem in financial modeling is to keep track of the price dynamics for those tradable assets. As for the fair price of these products, the crucial requirement is that we exclude arbitrages, which is defined as the following:

Definition 2.1. A trading strategy is a stochastic process $\{X_t\}$ that specifies the operations on the financial market on every time t. A strategy is called an arbitrage if it satisfies the following:

$$X_0 = 0,\ \mathbb{P}(X_T \geq 0) = 1,\ \mathbb{P}(X_T > 0) > 0$$

for some terminal time T.

An easy way to check whether there is arbitrage is to check the existence of the so-called risk-neutral measure:

Definition 2.2. A probability measure $\widetilde{\mathbb{P}}$ is said to be risk-neutral with respect to the stocks $\{S_i\}$ and probability measure $\mathbb{P}$ if (i) $\widetilde{\mathbb{P}}$ and $\mathbb{P}$ are equivalent (i.e., for every $A \in \mathcal{F}, \mathbb{P}(A) = 0$ if and only if $\widetilde{\mathbb{P}}(A) = 0$), and (ii) under $\widetilde{\mathbb{P}}$, the discounted stock price $D(t)S_i(t)$ is a martingale for every $i = 1, \ldots, m$, where $D(t)$ is the discounted factor.

In other words, it is to say that the expected return of holding any asset is the same under this new probability measure $\widetilde{\mathbb{P}}$. The relationship between the existence of the risk-neutral measure and the existence of arbitrage is provided by the following theorem, which we are going to prove later:

Theorem 2.3. (First fundamental theorem of asset pricing) If a market model has a risk-neutral probability measure, then it does not admit arbitrage.

2.2 One-period model

Suppose we have m assets in a financial market with initial prices $p = (p_1, p_2, \cdots, p_m)$, and at terminal time the market state has n possibilities, in which the asset price of the i-th asset at state j is given by X_{ij} for some $n \times m$ matrix X. In this model, a trading strategy can be simplified as specifying the amount of assets that the investor trade at the initial time and sell all assets at terminal time. As a result, the existence of arbitrage in this model can be formulated as follows:

Theorem 2.4. In the model we set up above, there is arbitrage iff there exists a vector $a \in \mathbb{R}^m$ s.t. $a^\top p = 0$, $X^\top a \geq 0$ and $X^\top a$ is a nonzero vector.

By Farkas-Stiemke Lemma in optimization theory, it is equivalent to:

Theorem 2.5. In the model we set up above, there is arbitrage iff there exists a vector $y \in \mathbb{R}^n$ that is positive at every component s.t. $p = Xy$.

By applying this theorem, we can construct a risk-neutral probability measure, which justifies the first fundamental theorem of asset pricing:

Theorem 2.6. Under the no arbitrage condition, suppose $y = (y_1, y_2, \cdots, y_n)^\top$, and the probability for the market to be in state j is $\tilde{y_j} = \frac{y_j}{\sum_{i=1}^m y_i}$, then this probability measure is a risk-neutral probability measure. Furthermore, if this market contains a risk-free asset with interest rate r, then the price for the assets p_i can be represented as $p_i = \frac{1}{1+r}\widetilde{\mathbb{E}}[X_i]$.

Proof. By no arbitrage condition, if we set $D = \sum_{j=1}^n y_j$,$p = Xy$ implies that $p_i = \sum_{j=1}^n X_{ij}y_j = D\sum_{j=1}^n X_{ij}\tilde{y_j} = D\widetilde{\mathbb{E}}[X_i]$. Hence for each asset i, the expected return rate is $\frac{E[X_i]-p_i}{p_i} = \frac{1}{D} - 1$, which does not vary with the type of asset i. Then the probability measure we set up is a risk-neutral measure. $\square$

Now suppose we have a new financial product in this market with return V_j at state j of the market, then we can price this product by this risk-neutral probability measure: the fair price is given by $u = D\widetilde{\mathbb{E}}[V] = \sum_{j=1}^n V_j y_j$. As a result, if the risk-neutral probability $\widetilde{\mathbb{P}}$ is not unique, the fair price of this new product would not be unique and usually becomes an interval. This gives rise to the completeness of the market:

Definition 2.7. (Completeness of the market) If for any r.v. V with possible values $v_1, v_2, \ldots, v_n$ s.t. $\exists a \in \mathbb{R}^n, v_j = \sum_{j=1}^m a_i X_{ij}$, where X is given above, then this market is complete.

Remark. Intuitively, it says that any new financial product can be replicated.

The relation between completeness and the uniqueness of $\widetilde{\mathbb{P}}$ is stated as follows:

Theorem 2.8. Uniqueness of $\widetilde{\mathbb{P}}$ is equivalent to completeness of the market.

Actually, for the system $p = Xy$:

(1) If $m > n$: there may not be solutions, and hence some products cannot be hedged $\Rightarrow$ arbitrage opportunity.

(2) If $m < n$: there may be many solutions, some uncertainty cannot be hedged $\Rightarrow$ incompleteness of the market. However, in practice the no-arbitrage condition is more important than the no-arbitrage condition.

Remark. We will focus on continuous time models, which often admit close-form solutions. In practice, people also use discrete time models (like GARCH-type models). However, discrete-time models capture the feature of financial derivatives more easily and are easier to deal with.

2.3 Diffusive market model

We built up this model based on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a filtration $\{\mathcal{F}_t\}$ with the following assumptions:

(1) We have a risk-free asset with stochastic interest rate r_t (which is adapted to $\{\mathcal{F}_t\}$). As a result, the discounting factor $D(t) = e^{-\int_0^t r_s ds}$

(2) The m risky assets $S_t^{(i)}$ are driven by the Brownian motion $\left(w_t^{(1)}, \cdots, w_t^{(d)}\right)$:

$$dS_t^{(i)} = \alpha_+^{(i)} S_t^{(i)} dt + S_t^{(i)} \sum_{j=1}^d \sigma_t^{(ij)} dw_t^{(j)}$$

The expected return rate for asset i is $E\left[\frac{dS_t^{(i)}}{S_t}\right] = \alpha_t^{(i)}$, and the volatility is

$$\begin{aligned}
&\operatorname{Var}\left[\frac{dS_t^{(i)}}{St}\right] \\
&= \operatorname{Var}\left[\sum_{j=1}^d \sigma_t^{(ij)} dW_t^{(j)}\right] \\
&= \sum_{j=1}^d \sigma_t^{(ij)2} \operatorname{Var}\left[dW_t^{(j)}\right] = \sum_{j=1}^d \sigma_t^{(ij)2} dt
\end{aligned}$$

The risky assets can be correlated:

$$\begin{aligned}
&\operatorname{Cov}\left(\frac{dS_t^{(i)}}{S_t^{(i)}}, \frac{dS_T^{(j)}}{S_T^{(j)}}\right) \\
&= \sum_{k=1}^{d}\sum_{l=1}^{d} \operatorname{Cov}\left(\sigma_t^{(ik)} dW_t^{(k)}, \sigma_t^{(jl)} dW_t^{(l)}\right) \\
&= \sum_{k=1}^{d}\sum_{l=1}^{(d)} \sigma_t^{(ik)} \sigma_t^{(jl)} \operatorname{Cov}\left(dW_t^{(k)}, dW_t^{(l)}\right) \\
&= \sum_{k=1}^{d} \sigma_t^(ik) \sigma_t^{(jk)} dt
\end{aligned}$$

$$\rho\left(\frac{dS_t^{(i)}}{S_t^{(i)}}, \frac{dS_t^{(j)}}{S_t^{(j)}}\right) = \frac{\sum_{k=1}^{d} \sigma_t^{(ik)} \sigma_t^{(jk)}}{\sqrt{\left(\sum_{k=1}^{d} \sigma_t^{(ik)2}\right)\left(\sum_{k=1}^{d} \sigma_t^{(ji)2}\right)}}.$$

If we denote $\sigma_t^{(i)} = \sum_{j=1}^{d} \sigma_t^{(ij)2}$ as the variance of the i-th asset, we can rewrite the SDE as

$$dS_t^{(i)} = \alpha_t^{(i)} S_t^{(i)} dt + \sigma_t^{(i)} S_t^{(i)} \sum_{j=1}^{d} \frac{\sigma_t^{(ij)}}{\sigma_t^{(i)}} dw_t^{(i)}$$

by setting $B_t^{(i)} = \sum_{j=1}^{d} \frac{\sigma_t^{(ij)}}{\sigma_t^{(i)}} dW_t^{(i)}$, which a BM (by checking Levy's characterization of Brownian motions), we have

$$dS_t^{(i)} = \alpha + S_t^{(i)} dt + \sigma_t^{(i)} S_t^{(i)} dB_t^{(i)}$$

We further relate it to the risk-free asset

$$\begin{aligned}
dS_t^{(i)} &= S_t^{(i)}\left(\alpha_t^{(i)} - r_t + r_t\right) dt + \sigma_t^{(i)} S_t^{(i)} dB_t^{(i)} \\
&= S_t^{(i)}\left(\lambda(i)\sigma_t^{(i)} + r_t\right) dt + \sigma_t^{(i)} S_t^{(i)} dB_t^{(i)}
\end{aligned}$$

where $\lambda_t^{(i)} = \frac{\alpha_t^{(i)} - r_t}{\sigma_t^{(i)}}$ is the extra return rate for unit volatility.

The dynamics of this discounted asset price can be calculated as:

$$\begin{aligned}
&d\left[D_t\right] = -r_t D_t dt \\
&d\left[D_t S_t^{(i)}\right] = D_t dS_t^{(i)} + S_t^{(i)} dD_t + dS_t^{(i)} dD_t \\
&= \left(\alpha_t^{(i)} - r_t\right) D_t S_t^{(i)} dt + D_t S_t^{(i)} \cdot \sum_{j=1}^{d} \sigma_t^{(ij)} dW_t^{(i)}
\end{aligned}$$

Now we consider trading strategies in this diffusive market model, which includes a risk-free asset, some risky assets and consumption. In this model, we assume that the assets have unlimited liquidity: we can buy or sell any amount of these assets without any transaction costs. The main task for trading strategies is to implement continuous rebalancing of the investment portfolio so that the payoff is optimized.

The trading strategies in the diffusive market model can be formulated as follows:

Definition 2.9. Let $\{X_t\}$ denotes the current wealth at time t, $\Delta_t^{(i)}$ denotes the amount of the asset i that the investor hold at time t, which is an adapted process with respect to $\{\mathcal{F}_t\}$ as we can only use current information for this continuous rebalancing, c_t be the consumption rate at time t. Then by specifying $\Delta_t^{(i)}$, c_t, a trading strategy is built and the dynamics of the wealth is given by

$$dX_t = r_t\left[X_t - \sum_{i=1}^{m} \Delta_t^{(i)} S_t^{(i)}\right] dt + \sum_{i=1}^{m} \Delta_t^{(i)} dS_t^{(i)} - c_t dt$$

If we do not allow consumption ($c_t > 0$) or injection ($c_t < 0$) of wealth, then the wealth process $\{X_t\}$ becomes a self-financing trading strategy when c_t is identically zero:

Definition 2.10. (Self-financing trading strategy) The wealth process $\{X_t\}$ is called a self-financing process if $c_t = 0$

in the formula above, whose dynamics under differential form is given by

$$\begin{aligned} dX_t &= r_t \left[X_t - \sum_{i=1}^{m} \Delta_t^{(i)} S_t^{(i)} \right] dt + \sum_{i=1}^{m} \Delta_t^{(i)} dS_t^{(i)} \\ &= r_t X_t dt + \sum_{i=1}^{m} \Delta_t^{(i)} \left[dS_t^{(i)} - r_t S_t^{(i)} dt \right] \end{aligned}$$

Furthermore, the dynamics of the discounted wealth process is given by

$$\begin{aligned} d\left[D_t X_t\right] &= D_t dX_t + X_t dD_t + dD_t dX_t \\ &= -r_t D_t X_t dt - r_t D_t dt \left(r_t X_t dt + \sum_{i=1}^{m} \Delta_t^{(i)} \left[dS_t^{(i)} - r_t S_t^{(i)} dt \right] \right) + D_t \left(r_t X_t dt + \sum_{i=1}^{m} \Delta_t^{(i)} \left[dS_t^{(i)} - r_t S_t^{(i)} dt \right] \right) \\ &= -r_t D_t X_t dt + r_t D_t X_t dt + \sum_{i=1}^{m} \Delta_t^{(i)} \left[D_t dS_t^{(i)} - r_t D_t S_t^{(i)} dt \right] \\ &= \sum_{i=1}^{m} \Delta_t^{(i)} d \left[D_t S_t^{(i)} \right] \end{aligned}$$

Of course, we would like to investigate the existence of arbitrage in this model. First we recall the definition of arbitrage:

Definition 2.11. Suppose X_t is an SFTS s.t. $X_0 = 0$, and there exists $T > 0$ s.t.

$$\mathbb{P}\left(X_T \geq 0\right) = 1, \quad P\left(X_T > 0\right) > 0$$

then this SFTS is an arbitrage.

To analyze the existence of arbitrage, we need the following definition of risk-neutral probability measure:

Definition 2.12. A probability measure $\widetilde{\mathbb{P}}$ is called a risk-neutral probability measure (with respect to the original probability measure $\mathbb{P}$) if $\widetilde{\mathbb{P}}$ is equivalent to $\mathbb{P}$ (which means that they agree on the events with zero/one/positive probability) and under $\widetilde{\mathbb{P}}$, $D_t S_t^{(i)}$ is a martingale for all $1 \leq i \leq n$.

Now we prove the first fundamental theorem of asset pricing under this model:

Theorem 2.13. (First fundamental theorem of asset pricing) If there is a risk neutral probability measure under this market model, then there is no arbitrage under this model.

Proof. Suppose $\widetilde{\{\mathbb{P}\}}$ is a risk-neutral probability measure and there is an arbitrage opportunity $\{X_t\}$:

$$X_0 = 0, \mathbb{P}(X_T \geq 0) = 1, \mathbb{P}(X_T > 0) > 0$$

Note that these relations also holds for $\widetilde{\mathbb{P}}$, by taking expectations, we have

$$\widetilde{\mathbb{P}}[D_T X_T > 0] > 0, \widetilde{\mathbb{E}}[D_T X_T] > 0 = \widetilde{\mathbb{E}}[D_0 X_0]$$

As $D_t S_t^{(i)}$ is a martingale under $\widetilde{\mathbb{P}}$ and we have

$$d[D_t X_t] = \sum_{i=1}^{m} \Delta_t^{(i)} d[D_t S_t^{(i)}]$$

which implies that $D_t X_t$ is a martingale, contradicting the inequality above. □

Furthermore, we want to find a specific risk-neutral probability measure, which is the task of Girsanov's theorem. If we would like the new probability measure be of the form in the theorem, i.e., let $\{Z_t\}_{t\geqslant 0}^{T}$ be a stochastic process satisfying

$$Z_t = \exp\left(-\int_0^t \theta_u dW_u - \frac{1}{2}\int_0^t \theta_u^2 du\right) \quad \forall 0 \leq t \leq T$$

and $Z_T = \frac{d\widetilde{\mathbb{P}}(\omega)}{d\mathbb{P}(\omega)}$ for a new probability measure $\widetilde{\mathbb{P}}$. Then the dynamics of $D_tS_t^{(i)}$ is given by

$$\begin{aligned} d\left[D_tS_t^{(i)}\right] &= D_tS_t^{(i)}\left[\left(\alpha_t^{(i)} - r_t\right)dt + \sum_{j=1}^{d}\sigma_t^{(ij)}\left[d\widetilde{W}_t^{(j)} - \theta_t^{(j)}dt\right]\right] \\ &= D_tS_t^{(i)}\left[\left(\alpha_t^{(i)} - r_t - \sum_{j=1}^{d}\sigma_t^{(ij)}\theta_t^{(ij)}\right)dt + \sum_{j=1}^{d}\sigma_t^{(ij)}d\widetilde{W}_t^{(j)}\right] \end{aligned}$$

As it is a martingale, we require that the drift term to be zero, which gives for any $i \in \{1, 2, \cdots, m\}$

$$\alpha_t^{(i)} - r_t = \sum_{j=1}^{d}\sigma_t^{(ij)}\theta_t^{(i)}$$

or in vector form,

$$\alpha_t - r_t 1 = \sigma_t\theta_t$$

where α_t, $\sigma_t = \left(\sigma_t^{(ij)}\right)$, θ_t are vectors with the i^{th} entry being the corresponding parameter.

Then for any risky asset $S_t^{(i)}$ under $\widetilde{\mathbb{P}}$, as the process defined by $\widetilde{W}_t = W_t + \int_0^t \theta_u du$ is a Brownian Motion w.r.t. $\tilde{P}$, its dynamics is given by

$$\begin{aligned} dS_t^{(i)} &= \alpha_t^{(i)}S_t^{(i)}dt + S_t^{(i)}\sum_{j=1}^{d}\sigma_t^{(ij)}\left(d\widetilde{W}_t^{(i)} - \theta_t^{(i)}dt\right) \\ &= \left(\alpha_t^{(i)} - \sum_{j=1}^{d}\sigma_t^{(ij)}\theta_t^{(i)}\right)S_t^{(i)} + S_t^{(i)}\sum_{j=1}^{d}\sigma_t^{(j)}d\widetilde{W}_t^{(i)} \\ &= r_tS_t^{(i)}dt + S_t^{(i)}\sum_{j=1}^{d}\sigma_t^{(ij)}d\widetilde{W}_t^{(j)} \end{aligned}$$

which implies that its expected return rate is the same as the risk-free asset.

Similar to the one-period model, we also would like to investigate the completeness of the market model:

Definition 2.14. For any $\mathcal{F}_T$-measurable r.v. V_T. there is a SFTS $\{X_t\}$ such that $X_T = V_T$ then this market is called complete.

In other words, in a complete market, any new risky asset can be replicated by an SFTS in the model.

An important tool in investigating the completeness of the market model is the following martingale representation theorem:

Theorem 2.15. (Martingale representation theorem) Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with d-dimensional Brownian motion $W_t = \left(W_t^{(1)}, \cdots, W_t^{(d)}\right)$. Then for any martingale $\{M_t\}$ w.r.t. $\{\mathcal{F}_t\}$, there is an d-dimensional adapted process $\Gamma_t = \left(\Gamma_t^{(1)}, \cdots, \Gamma_t^{(d)}\right)$ s.t.

$$M_t = M_0 + \sum_{j=1}^{d}\int_0^t \Gamma_u^{(j)}dW_u^{(j)}$$

Suppose we would like to replicate a new risky asset V_t with terminal payoff V_T. Assume that there is no arbitrage opportunity (as otherwise we can 'generate' wealth from nothing), then there exists a risk-neutral probability measure $\widetilde{\mathbb{P}}$ s.t. $D_tV_t = \widetilde{\mathbb{E}}\left[D_TV_T \mid \mathcal{F}_t\right]$. To replicate V_t by an SFTS X_t, it suffices to specify $\Delta_t^{(i)}$ for $i = 1, 2, \cdots, m$. By calculation before, the dynamics of D_tX_t is given by

$$d\left[D_tX_t\right] = \sum_{j=1}^{d}\sum_{i=1}^{m}\Delta_t^{(i)}D_tS_t^{(i)}\sigma_t^{(ij)}\widetilde{W}_t^{(j)}$$

As D_tV_t is a martingale under $\widetilde{\mathbb{P}}$, by martingale representation theorem, there exists an adapted process $\Gamma_t = \left(\Gamma_t^{(1)}, \cdots, \Gamma_t^{(d)}\right)$ s.t.

$$d\left[D_tV_t\right] = \sum_{j=1}^{d}P_t^{(j)}\tilde{w}_t^{(j)}$$

So it is sufficient that $X_0 = V_0$ and $\forall j = 1, 2, \cdots, d$,

$$\Gamma_t^{(j)} = \sum_{i=1}^{m}\Delta_t^{(i)}D_tS_t^{(i)}\sigma_t^{(ij)}$$

or in vector form,

$$\Gamma_t = \sigma_t^\top \begin{pmatrix} \Delta_t^{(1)} D_t s_t^{(1)} \\ \vdots \\ \Delta_t^{(n)} D_t s_t^{(n)} \end{pmatrix}$$

By result in linear algebra, the system above has a solution for any $\Gamma_t \in \mathbb{R}^m \Leftrightarrow \sigma$ has full rank and $m \geqslant d \Leftrightarrow$ The system for existence of risk-neutral probability measure has a unique solution. Then we have the second fundamental theorem of asset pricing:

Definition 2.16. (Second fundamental theorem of asset pricing) The completeness of the market is equivalent to the existence of risk-neutral probability measure.

Remark. Completeness is not an essential requirement to the market model, but existence of risk-neutral probability measure is a nice property.

2.4 Assets with dividends

The diffusive market models contain only assets that does not pay dividends, which is different from real-life experience. In this section we would consider models that pay dividends continuously or at discrete times.

Definition 2.17. (Assets with continuously paying dividends) A stock that pays dividends continuously over time at a rate $A(t)$ per unit time, where $A(t)$, $0 \le t \le T$, is a nonnegative adapted process, has dynamics

$$dS(t) = \alpha(t)S(t)dt + \sigma(t)S(t)dW(t) - A(t)S(t)dt$$

An agent who holds the stock receives both the capital gain or loss due to stock price movements and the continuously paying dividend. Thus, if $\Delta(t)$ is the number of shares held at time t, then the portfolio value $X(t)$ satisfies

$$\begin{aligned} dX(t) &= \Delta(t)dS(t) + \Delta(t)A(t)S(t)dt + R(t)[X(t) - \Delta(t)S(t)]dt \\ &= R(t)X(t)dt + (\alpha(t) - R(t))\Delta(t)S(t)dt + \sigma(t)\Delta(t)S(t)dW(t) \\ &= R(t)X(t)dt + \Delta(t)S(t)\sigma(t)[\Theta(t)dt + dW(t)] \end{aligned}$$

where

$$\Theta(t) = \frac{\alpha(t) - R(t)}{\sigma(t)}$$

is the usual market price of risk. We define

$$\widetilde{W}(t) = W(t) + \int_0^t \Theta(u)du$$

and use Girsanov's Theorem to change to a measure $\widetilde{\mathbb{P}}$ under which $\widetilde{W}$ is a Brownian motion, so we may rewrite it as

$$dX(t) = R(t)X(t)dt + \Delta(t)S(t)\sigma(t)d\widetilde{W}(t)$$

The discounted portfolio value satisfies

$$d[D(t)X(t)] = \Delta(t)D(t)S(t)\sigma(t)d\widetilde{W}(t).$$

In particular, under the risk-neutral measure $\widetilde{\mathbb{P}}$, the discounted portfolio process is a martingale. Here we denote by $D(t) = e^{-\int_0^t R(u)du}$ the usual discount process.

If we now wish to hedge a short position in a derivative security paying $V(T)$ at time T, where $V(T)$ is an $\mathcal{F}(T)$-measurable random variable, we will need to choose the initial capital $X(0)$ and the portfolio process $\Delta(t)$, $0 \le t \le T$, so that $X(T) = V(T)$. Because $D(t)X(t)$ is a martingale under $\widetilde{\mathbb{P}}$, we must have

$$D(t)X(t) = \widetilde{\mathbb{E}}[D(T)V(T) \mid \mathcal{F}(t)], 0 \le t \le T.$$

The value $X(t)$ of this portfolio at each time t is the value (price) of the derivative security at that time, which we denote by $V(t)$. Making this replacement in the formula above, we obtain the risk-neutral pricing formula

$$D(t)V(t) = \widetilde{\mathbb{E}}[D(T)V(T) \mid \mathcal{F}(t)], 0 \le t \le T.$$

We have obtained the same risk-neutral pricing formula as in the case of no dividends. Furthermore, conditions that guarantee that a short position can be hedged, and hence risk-neutral pricing is fully justified.

The difference between the dividend and no-dividend cases is in the evolution of the underlying stock under the risk-neutral measure. From the dynamics of the risky asset and the definition of $\widetilde{W}(t)$, we see that

$$dS(t) = [R(t) - A(t)]S(t)dt + \sigma(t)S(t)d\widetilde{W}(t)$$

Under the risk-neutral measure, the stock does not have mean rate of return $R(t)$, and consequently the discounted stock price is not a martingale. Indeed,

$$S(t)=S(0)\exp\left\{\int_0^t \sigma(u)d\widetilde{W}(u)+\int_0^t\left[R(u)-A(u)-\frac{1}{2}\sigma^2(u)\right]du\right\}.$$

The process

$$e^{\int_0^t A(u)}D(t)S(t)=\exp\left\{\int_0^t \sigma(u)d\widetilde{W}(u)-\frac{1}{2}\int_0^t\sigma^2(u)du\right\}$$

is a martingale. This is the interest-rate-discounted value at time t of an account that initially purchases one share of the stock and continuously reinvests the dividends in the stock.

Finally, let us consider the case when the dividend is paid in lumps. That is to say there are times $0<t_1<t_2<t_n<T$ and, at each time t_j, the dividend paid is $a_jS\left(t_j-\right)$, where $S\left(t_j-\right)$ denotes the stock price just prior to the dividend payment. The stock price after the dividend payment is the stock price before the dividend payment less the dividend payment:

$$S\left(t_j\right)=S\left(t_j-\right)-a_jS\left(t_j-\right)=\left(1-a_j\right)S\left(t_j-\right).$$

We assume that each a_j is an $\mathcal{F}\left(t_j\right)$-measurable random variable taking values in $[0,1]$. If $a_j=0$, no dividend is paid at time t_j. If $a_j=1$, the full value of the stock is paid as a dividend at time t_j, and the stock value is zero thereafter. To simplify the notation, we set $t_0=0$ and $t_{n+1}=T$. However, neither $t_0=0$ nor $t_{n+1}=T$ is a dividend payment date (i.e., $a_0=0$ and $a_{n+1}=0$). We assume that, between dividend payment dates, the stock price follows a generalized geometric Brownian motion:

$$dS(t)=\alpha(t)S(t)dt+\sigma(t)S(t)dW(t),t_j\leq t<t_{j+1},j=0,1,\ldots,n$$

Between dividend payment dates, the differential of the portfolio value corresponding to a portfolio process $\Delta(t),0\leq t\leq T$, is

$$\begin{aligned}dX(t)&=\Delta(t)dS(t)+R(t)[X(t)-\Delta(t)S(t)]dt\\&=R(t)X(t)dt+(\alpha(t)-R(t))\Delta(t)S(t)dt+\sigma(t)\Delta(t)S(t)dW(t)\\&=R(t)X(t)dt+\Delta(t)\sigma(t)S(t)[\Theta(t)dt+dW(t)],\end{aligned}$$

where the market price of risk $\theta(t)$ is again defined as above. At the dividend payment dates, the value of the portfolio stock holdings drops by $a_j\Delta\left(t_j\right)S\left(t_j-\right)$, but the portfolio collects the dividend $a_j\Delta\left(t_j\right)S\left(t_j-\right)$, and so the portfolio value does not jump. It follows that

$$dX(t)=R(t)X(t)dt+\Delta(t)\sigma(t)S(t)[\Theta(t)dt+dW(t)]$$

is the correct formula for the evolution of the portfolio value at all times t. We again define $\widetilde{W}$ and change to a measure $\widetilde{\mathbb{P}}$ under which $\widetilde{W}$ is a Brownian motion, and obtain the risk-neutral pricing formula same as above.

2.5 Derivatives pricing: forwards and futures

After setting up models for risky assets, we can analyze the value of financial derivatives written on these risky assets. We first give a detail definition of a forward:

Definition 2.18. A forward contract is an agreement to pay a specified delivery price K at a delivery date T, for the asset whose price at time t is $S(t)$. The T-forward price $\mathrm{For}_S(t,T)$ of this asset at time t, where $0\leq t\leq T$, is the value of K that makes the forward contract have no-arbitrage price zero at time t.

Then we consider the market model with risky assets introduced in the previous section with a risk-neutral probability measure and a zero-coupon bond that we can trade with price $B(t,T)=\mathbb{E}\left[\frac{D_T}{D_t}\mid\mathcal{F}_t\right]$ at time t, where T is the maturity. We first give the fair price for forwards:

Theorem 2.19. Assume that zero-coupon bonds of all maturities can be traded. Then

$$\mathrm{For}_S(t,T)=\frac{S(t)}{B(t,T)},\quad 0\leq t\leq T$$

Proof. Suppose that at time t an agent sells the forward contract with delivery date T and delivery price K. Suppose further that the value K is chosen so that the forward contract has price zero at time t. Then selling the forward contract generates no income. Having sold the forward contract at time t, suppose the agent immediately shorts $\frac{S(t)}{B(t,T)}$ zero-coupon bonds and uses the income $S(t)$ generated to buy one share of the asset. The agent then does no further trading until time T, at which time she owns one share of the asset, which she delivers according to the forward contract. In exchange, she receives K. After covering the short bond position, she is left with $K-\frac{S(t)}{B(t,T)}$. If this is positive, the agent has found an arbitrage. If it is negative, the agent could instead have taken the opposite position, going long the forward, long the T-maturity bond, and short the asset, to again achieve an arbitrage. In order to preclude arbitrage, K must be given by the formula above. $\square$

We can also give a proof in the view of risk-neutral probability measure:

Proof. Because we have assumed the existence of a risk-neutral measure and are pricing all assets by the risk-neutral pricing formula, we must be able to obtain the formula above from the risk-neutral pricing formula as well. Indeed, using the fact that the discounted asset price is a martingale under $\widetilde{\mathbb{P}}$, we compute the price at time t of the forward contract to be

$$\begin{aligned}&\frac{1}{D(t)}\widetilde{\mathbb{E}}[D(T)(S(T)-K)\mid\mathcal{F}_t]\\&=\frac{1}{D(t)}\widetilde{\mathbb{E}}[D(T)S(T)\mid\mathcal{F}_t]-\frac{K}{D(t)}\widetilde{\mathbb{E}}[D(T)\mid\mathcal{F}_t]\\&=S(t)-KB(t,T).\end{aligned}$$

In order for this to be zero, K must be given by the formula above. $\square$

Remark. The forward price, as shown above, is irrelevant to the dynamics of the underlying asset price.

However, forwards have certain drawbacks:

(a) They are traded over-the-counter but not in the exchange, as a result, they cannot be traded. (b) There is possibility that $S_T > \mathrm{For}_S(t,T)$. Under this situation, the short position may suffer from loss and have intention to default.

To resolve such drawbacks, we introduce a new financial derivative called future.

Definition 2.20. A future contract an agreement to pay a specified delivery price K at a delivery date T, for the asset whose price at time t is $S(t)$. The T-future price $\mathrm{Fut}_S(t,T)$ of this asset at time t, where $0\le t\le T$, is the value of K that makes the future contract have no-arbitrage price zero at time t. However, the main difference between future and forward is that futures can be traded in the exchange market.

To price the future contract, consider a time interval $[0,T]$, which we divide into subintervals using the partition points $0=t_0<t_1<t_2<\cdots<t_n=T$. We shall refer to each subinterval $[t_k,t_{k+1})$ as a "day". Then if the future price goes up in a day, it benefits the short position, otherwise it benefits the long position.

Suppose now we only consider the long position, at time t_i the long position would receive cash flow $\mathrm{Fut}(t_i,T)-\mathrm{Fut}(t_{i-1},T)$ in the margin account.

We would like to calculate the fair price $\mathrm{Fut}(t,T)$. To do this, we first stand at time t_i and then at this time the future price is $\mathrm{Fut}(t_i,T)$. When we go to time t_{i+1} we receive a cash flow $\mathrm{Fut}(t_{i+1},T)-\mathrm{Fut}(t_i,T)$. By no-arbitrage condition we have

$$\widetilde{\mathbb{E}}\left[\frac{D_{t_{i+1}}}{D_{t_i}}(\mathrm{Fut}(t_{i+1})-\mathrm{Fut}(t_i,T))\mid\mathcal{F}_{t_i}\right]$$

If we assume that $t_{i+1}-t_i$ is very small (or the interest rate is constant throughout the day), then $\frac{D_{t_{i+1}}}{D_{t_i}}$ can be regarded as $\mathcal{F}_{t_i}$-measurable, and then we have

$$\widetilde{\mathbb{E}}\left[\mathrm{Fut}(t_{i+1})\mid\mathcal{F}_{t_i}\right]=\mathrm{Fut}(t_i,T)$$

As we have the terminal condition $\mathrm{Fut}(T,T)=S_T$, inductively we can have

$$\mathrm{Fut}(t,T)=\widetilde{\mathbb{E}}\left[S_T\mid\mathcal{F}_t\right]$$

We can also compute the price of a futures contract in the view of risk-neutral probability measure:

Theorem 2.21. The fair price of a futures price $\mathrm{Fut}(t,T)$ at time t with maturity T with underlying asset S_t is given by $\widetilde{\mathbb{E}}\left[S_T\mid\mathcal{F}_t\right]$

Proof. Suppose the SFTS $\{X_t\}$ holds Δ_t shares of underlying asset and Γ_t shares of futures contract. Then as the futures contract does not cause any consumption of the SFTS but only results in a wealth income or loss, the dynamics of the SFTS is given by

$$dX_t=r_t(X_t-\Delta_tS_t)dt+\Delta_tdS_t+\Gamma_td\,\mathrm{Fut}(t,T)$$

Then a simple calculation shows that the discounted wealth process has dynamics

$$d\left[D_tX_t\right]=\Delta_td\left[D_tS_t\right]+D_t\Gamma_td\,\mathrm{Fut}(t,T)$$

To ensure that it is still a martingale under the risk-neutral probability measure, it is sufficient that $\mathrm{Fut}(t,T)$ is a martingale under $\widetilde{\mathbb{P}}$. And as we have $\mathrm{Fut}(T,T)=S_T$, the fair price of $\mathrm{Fut}(t,T)$ is given by $\widetilde{\mathbb{E}}\left[S_T\mid\mathcal{F}_t\right]$. $\square$

Remark. We conclude with a comparison of forward and futures prices. We have defined these prices to be

$$\begin{aligned}\mathrm{For}_S(t,T) &= \frac{S(t)}{B(t,T)},\\ \mathrm{Fut}_S(t,T) &= \widetilde{\mathbb{E}}[S(T)\mid\mathcal{F}(t)].\end{aligned}$$

If the interest rate is a constant r, then $B(t,T)=e^{-r(T-t)}$ and

$$\begin{aligned}\mathrm{For}_S(t,T) &= e^{r(T-t)}S(t),\\ \mathrm{Fut}_S(t,T) &= e^{rT}\widetilde{\mathbb{E}}\left[e^{-rT}S(T)\mid\mathcal{F}(t)\right]=e^{rT}e^{-rt}S(t)=e^{r(T-t)}S(t).\end{aligned}$$

In this case, the forward and futures prices agree. We compare $\mathrm{For}_S(0,T)$ and $\mathrm{Fut}_S(0,T)$ in the case of a random interest rate. In this case, $B(0,T)=\widetilde{\mathbb{E}}D(T)$, and the so-called forward-futures spread is

$$\begin{aligned}\mathrm{For}_S(0,T)-\mathrm{Fut}_S(0,T) &= \frac{S(0)}{\widetilde{\mathbb{E}}D(T)}-\widetilde{\mathbb{E}}S(T)\\ &= \frac{1}{\widetilde{\mathbb{E}}D(T)}\{\widetilde{\mathbb{E}}[D(T)S(T)]-\widetilde{\mathbb{E}}D(T)\cdot\widetilde{\mathbb{E}}S(T)\}\\ &= \frac{1}{B(0,T)}\widetilde{\mathrm{Cov}}(D(T),S(T)),\end{aligned}$$

where $\widetilde{\mathrm{Cov}}(D(T),S(T))$ denotes the covariance of $D(T)$ and $S(T)$ under the risk-neutral measure. If the interest rate is nonrandom, this covariance is zero and the futures price agrees with the forward price.

2.6 Derivatives pricing: options

First we recall the definition of an option:

Definition 2.22. An option is a financial contract written on the underlying which specifies how to do the transaction of the underlying in the future. The party who sells the option has only obligation while the party who buys the option has only right of trading the asset.

The most classical type of option is the European options, which are defined as follows:

Definition 2.23. (European options) The European call (put) option is the financial contract that allows the option holder to buy (sell) the underlying asset at price K at time T, where K is called the option price and T is called the maturity.

As a result, the European call option has payoff $\max(S_T-K,0)$ and the European put option has payoff $\max(K-S_T,0)$. In general, the option payoff is a function of the price of the underlying asset S_T, denoted as $f(S_T)$, which is called the payoff function.

To price an option, we would consider a market model with a risk-free asset and the underlying asset together with the option having unlimited liquidity with a risk-neutral probability measure $\widetilde{\mathbb{P}}$. Similar as before, we can show that the fair price of an option is given by

$$\widetilde{\mathbb{E}}\left[\frac{D_T}{D_t}f(S_T)\mid\mathcal{F}_t\right]$$

where f is the payoff function. This is because the SFTS that holds only one share of the option only changes its value at the maturity with payoff $f(S_T)$, which should be the price of the option at this time. As discounted value of the SFTS should be a martingale under the risk-neutral probability measure, we obtain the formula for the fair price of the asset above.

2.7 The Black-Scholes formula

The Black-Scholes model is the diffusive market model having one risky asset with constant interest rate r, constant return rate μ, and volatility σ. Under this model, we can have many analytic results concerning the price of the risky assets and options in this market.

By result before , if we let $\frac{\mu-r}{\sigma}=\theta$, $\quad\frac{d\tilde{\mathbb{P}}(\omega)}{d\mathbb{P}(\omega)}=e^{-\frac{1}{2}\theta^2T-\theta W_T(\omega)}$, then $\tilde{\mathbb{P}}$ is the risky-neutral probability measure under which $e^{-rt}S_t$ is a martingale: the dynamics of S_t is given by

$$dS_t=rS_td_t+\sigma S_td\widetilde{W}_t$$

Now we want to derive an analytic formula for the options under the Black-Scholes model, which is formulated as follows:

Theorem 2.24. The European call option price at time t written on the underlying asset S_t is given by the Black-Scholes formula

$$C_t = S_t N(d_1) - Ke^{-r(T-t)} N(d_2)$$

where $d_1 = \frac{\ln(S_t/K)+\left(r+\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}$, $d_2 = d_1 - \sigma\sqrt{T-t}$ and N is the cdf for the standard normal distribution.

Proof. We calculate the option price directly:

$$\begin{aligned}
V_t &= \widetilde{\mathbb{E}}\left[e^{-r(T-t)} \max(S_T - K, 0)|\mathcal{F}_t\right] \\
&= \widetilde{\mathbb{E}}\left[e^{-r(T-t)}(S_T - K)1_{\{S_T \geq K\}}|\mathcal{F}_t\right] \\
&= \widetilde{\mathbb{E}}\left[e^{-r(T-t)} S_T 1_{\{S_T \geq K\}}|\mathcal{F}_t\right] - e^{-r(T-t)} K\widetilde{\mathbb{E}}\left[1_{\{S_T \geq K\}}|\mathcal{F}_t\right]
\end{aligned}$$

To simplify the calculation, consider change of measure driven by

$$\frac{d\hat{\mathbb{P}}_{(\omega)}}{d\widetilde{\mathbb{P}}(\omega)} = \frac{e^{-r(T-t))} S_T}{s_t} = e^{-\frac{1}{2}\sigma^2(T-t)+\sigma\left(\widetilde{W}_T - \widetilde{W}_t\right)}$$

By Girsanov's theorem, $\hat{W}_t = \widetilde{W}_t - \sigma t$ is a standard BM under $\hat{\mathbb{P}}$, and then we can calculate

$$\begin{aligned}
\widetilde{\mathbb{E}}\left[e^{-r(T-t)} S_T 1_{\{S_T \geq K\}}|\mathcal{F}_t\right] &= \widetilde{\mathbb{E}}\left[\frac{e^{-r(T-t)} S_T}{S_t} S_t 1_{\{S_T \geq K\}}|\mathcal{F}_t\right] \\
=& \hat{\mathbb{E}}\left[S_t 1_{\{S_T \geq K\}}|\mathcal{F}_t\right] = S_t \hat{\mathbb{E}}\left[1_{\{S_T \geq K\}}|\mathcal{F}_t\right]
\end{aligned}$$

as the dynamics of S_t is given by

$$dS_t = (r + \sigma^2) S_t dt + \sigma S_t d\hat{W}_t$$

we have

$$\hat{\mathbb{E}}\left[1_{\{S_T \geq K\}}|\mathcal{F}_t\right] = \hat{\mathbb{P}}\left[S_T \geq K|\mathcal{F}_t\right] = N(d_1)$$

Similarly,

$$\widetilde{\mathbb{E}}\left[1_{\{S_T \geq K\}}|\mathcal{F}_t\right] = \widetilde{\mathbb{P}}\left[S_T \geq K|\mathcal{F}_t\right] = N(d_2)$$

as the dynamics of S_t is given by

$$dS_t = rS_t d_t + \sigma S_t d\widetilde{W}_t$$

□

There is a simple relation between the European call and put options regardless of the model we are considering:

Theorem 2.25. (Put-call parity) For any model with constant return rate r, we have

$$C_t - P_t = S_t - e^{-r(T-t)} K$$

Proof. By putting things together, we have

$$\begin{aligned}
C_t - P_t &= \widetilde{\mathbb{E}}\left[e^{-r(T-t)} \max(S_T - K, 0)|\mathcal{F}_t\right] - \widetilde{\mathbb{E}}\left[e^{-r(T-t)} \max(K - S_T, 0)|\mathcal{F}_t\right] \\
&= \widetilde{\mathbb{E}}\left[e^{-r(T-t)}(\max(S_T - K, 0) - \max(K - S_T, 0))|\mathcal{F}_t\right] = \widetilde{\mathbb{E}}\left[e^{-r(T-t)}(S_T - K)|\mathcal{F}_t\right] \\
&= \widetilde{\mathbb{E}}\left[e^{-r(T-t)} S_T|\mathcal{F}_t\right] - Ke^{-r(T-t)} = S_t - e^{-r(T-t)} K
\end{aligned}$$

□

We can also compute the price of an European call option by the Feymann-Kac theory, which gives rise the Black-Scholes PDE:

Theorem 2.26. (Black-Scholes PDE) The call option price $C = c(t, S_t)$ under the Black-Scholes model satisfies the PDE

$$C_t + rxC_x + \frac{1}{2}\sigma^2 x^2 C_{xx} - rC = 0$$

with boundary condition $c(T, x) = \max(x - T, 0)$.

Proof. We begin by computing the differential of $c(t, S(t))$. By Ito formula, it is

$$
\begin{aligned}
&dc(t, S(t)) \\
=&c_t(t, S(t))dt + c_x(t, S(t))dS(t) + \frac{1}{2}c_{xx}(t, S(t))dS(t)dS(t) \\
=&c_t(t, S(t))dt + c_x(t, S(t))(\alpha S(t)dt + \sigma S(t)dW(t)) + \frac{1}{2}c_{xx}(t, S(t))\sigma^2 S^2(t)dt \\
=&\left[c_t(t, S(t)) + \alpha S(t)c_x(t, S(t)) + \frac{1}{2}\sigma^2 S^2(t)c_{xx}(t, S(t))\right]dt + \sigma S(t)c_x(t, S(t))dW(t)
\end{aligned}
$$

We next compute the differential of the discounted option price $e^{-rt}c(t, S(t))$. Let $f(t, x) = e^{-rt}x$. According to the Ito formula,

$$
\begin{aligned}
&d\left(e^{-rt}c(t, S(t))\right) \\
=&df(t, c(t, S(t))) \\
=&f_t(t, c(t, S(t)))dt + f_x(t, c(t, S(t)))dc(t, S(t)) + \frac{1}{2}f_{xxt}(t, c(t, S(t)))dc(t, S(t))dc(t, S(t)) \\
=&-re^{-rt}c(t, S(t))dt + e^{-rt}dc(t, S(t)) \\
=&e^{-rt}\left[-rc(t, S(t)) + c_t(t, S(t)) + \alpha S(t)c_x(t, S(t)) + \frac{1}{2}\sigma^2 S^2(t)c_{xx}(t, S(t))\right]dt + e^{-rt}\sigma S(t)c_x(t, S(t))dW(t)
\end{aligned}
$$

As it is should be martingale under the risk-neutral probability measure, the drift term should be zero and we have the desired result. □

After deriving this Black-Scholes PDE, we can hedge the call option by its delta: recall that for any SFTS $\{X_t\}$ holding Δ_t shares of stock, the dynamics of its discounted price is given by

$$d\left[e^{-rt}X_t\right] = e^{-rt}\sigma S_t \Delta_t d\widetilde{W}_t$$

So it is sufficient that

$$X_0 = c(0, S_0), \quad \Delta_t = c_x(t, S_t)$$

And the analytic formula of c_x can be computed as

$$
\begin{aligned}
c(t, S_t) &= \widetilde{\mathbb{E}}\left[e^{-r(T-t)}(S_T - K)1_{\{S_T \geq K\}}|\mathcal{F}_t\right] \\
\Rightarrow c_x(t, S_t) &= \widetilde{\mathbb{E}}\left[e^{-r(T-t)}e^{(r-\frac{1}{2}\sigma^2(T-t))+\sigma(\widetilde{W}_T-\widetilde{W}_t)}1_{\{S_T \geq K\}}|\mathcal{F}_t\right] \\
&= \widetilde{\mathbb{E}}\left[e^{-\frac{1}{2}\sigma^2(T-t)+\sigma(\widetilde{W}_T-\widetilde{W}_t)}1_{\{S_T \geq K\}}|\mathcal{F}_t\right] \\
&= \hat{\mathbb{E}}\left[1_{\{S_T \geq K\}}|\mathcal{F}_t\right] = N(d_1)
\end{aligned}
$$

The option Greeks are important in risk management, including:

- $\Delta_t = c_x(t, S_t)$: the option Delta
- $\Gamma_t = c_{xx}(t, S_t)$: the option Gamma
- $\Theta_t = c_t(t, S_t)$: the option Theta
- $\rho_t = \frac{\partial c(t, S_t)}{\partial r}$: the option Rho
- $U_t = \frac{\partial c(t, S_t)}{\partial \sigma}$: the option Vega.

2.8 Drawback of Black-Scholes model

In real life applications, the options are traded liquidly and their market price is affected by supply and demand. As a result, it is important to fit the model into real-world data, in other words, to do calibration.

The calibration of the Black-Scholes model can be formulated as follows:

As the dynamics of the risky asset is given by $dS_t = rS_t dt + \sigma S_t d\widetilde{W}_t$ under the risk-neutral probability measure $\widetilde{\mathbb{P}}$ and r is known, it suffices to optimize the volatility σ to fit the option prices. Moreover, under this model the option price is increasing w.r.t. σ (as we have calculated in homework 9), for any specific time T and strike price K, there must exist an implied volatility σ_{imp} such that the call option price under the Black-Scholes model equals its market price. However, if the option price does evolve as the Black-Scholes model describes, the implied volatility should be independent of time T and strike price K, and of course it cannot be the case. In real markets we often observe that the implied volatility depends on the strike price K and time T. The models that captures the dependence of implied volatility w.r.t. time are the term-structure models, which we shall introduce later. Here we first focus on the dependence of the implied volatility w.r.t. the strike price.

There are two kinds of dependence between them: the first one is that as the strike price K increase, σ_{imp} would decrease, which is called **the volatility smirk**; the second one is that the as K increase, σ_{imp} first decreases then increases, which is called **the volatility smile**. To capture such dependence, we need some new models to describe the dynamics of the underlying:

Definition 2.27. The local volatility model is the assumption that the dynamics of the risky asset S_t satisfies

$$dS_t = rS_t dt + \sigma(S_t)S_t d\widetilde{W}_t$$

In particular, when $\sigma(S_t) = \sigma_0 S_t^{\beta}$, it is called **constant-elasticity-of-variance (CEV model)**.

Remark. When $\beta < 0$ in the CEV model, it would result in a leverage effect and can capture the volatility smirk.

Definition 2.28. The stochastic volatility models are models with underlying asset price S_t s.t.

$$dS_t = rS_t dt + \sqrt{v_t} S_t d\widetilde{W}_t^{(1)}$$
$$dV_t = \alpha(V_t)dt + \beta(V_t)d\widetilde{W}_t^{(2)}$$

with $d\widetilde{W}_t^{(1)} d\widetilde{W}_t^{(2)} = \rho$.
In the Heston model, the dynamics of v_t is specified as

$$dv_t = \kappa(\theta - v_t)dt + \sigma\sqrt{v_t}d\widetilde{W}_t^{(2)}$$

When $\rho < 0$, the stochastic volatility model would result in a leverage effect.

3 Exotic options

The European calls and puts considered thus far in this text are sometimes called vanilla or even plain vanilla options. Their payoffs depend only on the final value of the underlying asset. Options whose payoffs depend on the path of the underlying asset or have limited liquidity are called path-dependent or exotic options.

3.1 Barrier options

There are several types of barrier options. Some "knock out" when the underlying asset price crosses a barrier (i.e., they become worthless). If the underlying asset price begins below the barrier and must cross above it to cause the knock-out, the option is said to be up-and-out. A down-and-out option has the barrier below the initial asset price and knocks out if the asset price falls below the barrier. Other options "knock in" at a barrier (i.e., they pay off zero unless they cross a barrier). Knock-in options also fall into two classes, up-and-in and down-and-in. The payoff at expiration for barrier options is typically either that of a put or a call. More complex barrier options require the asset price to not only cross a barrier but spend a certain amount of time across the barrier in order to knock in or knock out.

To price the up-and-out call options, one important quantity to analyze is the running maximum of a drifted Brownian motion, which can be done by combining the reflection principle and the Girsanov's theorem introduced in the first section.

Theorem 3.1. Let α be a given number, and define

$$\widehat{W}(t) = \alpha t + \widetilde{W}(t), \quad 0 \le t \le T.$$

This Brownian motion $\widehat{W}(t)$ has drift α under $\widetilde{\mathbb{P}}$. We further define

$$\widehat{M}(T) = \max_{0 \le t \le T} \widehat{W}(t)$$

Because $\widehat{W}(0) = 0$, we have $\widehat{M}(T) \ge 0$. We also have $\widehat{W}(T) \le \widehat{M}(T)$. Then the joint density under $\widetilde{\mathbb{P}}$ of the pair $(\widehat{M}(T), \widehat{W}(T))$ is

$$\widetilde{f}_{\widehat{M}(T),\widehat{W}(T)}(m, w) = \frac{2(2m - w)}{T\sqrt{2\pi T}} e^{\alpha w - \frac{1}{2}\alpha^2 T - \frac{1}{2T}(2m-w)^2}, \quad w \le m, m \ge 0,$$

and is zero for other values of m and w.

Proof. We define the exponential martingale

$$\widehat{Z}(t) = e^{-\alpha \widetilde{W}(t) - \frac{1}{2}\alpha^2 t} = e^{-\alpha \widehat{W}(t) + \frac{1}{2}\alpha^2 t}, \quad 0 \le t \le T,$$

and use $\widehat{Z}(T)$ to define a new probability measure $\widehat{\mathbb{P}}$ by

$$\widehat{\mathbb{P}}(A) = \int_A Z(T) d\widetilde{\mathbb{P}} \text{ for all } A \in \mathcal{F}.$$

By Girsanov's theorem, $\widehat{W}_t$ is a standard BM under $\widehat{\mathbb{P}}$. By reflection principle,

$$\widehat{f}_{\widehat{M}(T),\widehat{W}(T)}(m, w) = \frac{2(2m - w)}{T\sqrt{2\pi T}} e^{-\frac{1}{2T}(2m-w)^2}, \quad w \le m, m \ge 0$$

and is zero for other values of m and w. To work out the density of $(\widehat{M}(T), \widehat{W}(T))$ under $\widetilde{\mathbb{P}}$, the change of measure formula implies

$$\begin{aligned}
&\widetilde{\mathbb{P}}\{\widehat{M}(T) \le m, \widehat{W}(T) \le w\} \\
&\quad = \widetilde{\mathbb{E}}\left[\mathbb{I}_{\{\widehat{M}(T) \le m, \widehat{W}(T) \le w\}}\right] \\
&\quad = \widehat{\mathbb{E}}\left[\frac{1}{\widehat{Z}(T)} \mathbb{I}_{\{\widehat{M}(T) \le m, \widehat{W}(T) \le w\}}\right] \\
&\quad = \widehat{\mathbb{E}}\left[e^{\alpha \widehat{W}(T) - \frac{1}{2}\alpha^2 T_1} \mathbb{I}_{\{\widehat{M}(T) \le m, \widehat{W}(T) \le w\}}\right] \\
&\quad = \int_{-\infty}^{w} \int_{-\infty}^{m} e^{\alpha y - \frac{1}{2}\alpha^2 T} \widehat{f}_{\widehat{M}(T),\widehat{W}(T)}(x, y) dx dy.
\end{aligned}$$

Therefore, the density of $(\widehat{M}(T), \widehat{W}(T))$ under $\widetilde{\mathbb{P}}$ is

$$\frac{\partial^2}{\partial m \partial w} \widetilde{P}\{\widehat{M}(T) \le m, \widehat{W}(T) \le w\} = e^{\alpha w - \frac{1}{2}\alpha^2 T} \widehat{f}_{\widehat{M}(T),\widehat{W}(T)}(m, w)$$

$\square$

Corollary 3.2. By integration we also have

$$\widetilde{\mathbb{P}}\{\widehat{M}(T) \leq m\} = N\left(\frac{m-\alpha T}{\sqrt{T}}\right) - e^{2\alpha m} N\left(\frac{-m-\alpha T}{\sqrt{T}}\right), \quad m \geq 0$$

and the density under $\widetilde{\mathbb{P}}$ of the random variable $\widehat{M}(T)$ is

$$\tilde{f}_{\widehat{M}(T)}(m) = \frac{2}{\sqrt{2\pi T}} e^{-\frac{1}{2T}(m-\alpha T)^2} - 2\alpha e^{2\alpha m} N\left(\frac{-m-\alpha T}{\sqrt{T}}\right), m \geq 0, \ (7.2.7)$$

and is zero for $m < 0$.

After having this, we can derive an analytic formula for the option price:

The risk-neutral price at time zero of the up-and-out call with payoff $V(T)$ is $V(0) = \widetilde{\mathbb{E}}\left[e^{-rT}V(T)\right]$ and $V(T)$ is given by

$$\begin{aligned}
V(T) &= \left(S(0)e^{\sigma \widehat{W}(T)} - K\right)^+ \mathbb{I}_{\left\{S(0)e^{\sigma \widehat{M}(T)} \leq B\right\}} \\
&= \left(S(0)e^{\sigma \widehat{W}(T)} - K\right) \mathbb{I}_{\left\{S(0)e^{\sigma \widehat{W}(T)} \geq K, S(0)e^{\sigma \widehat{M}(T)} \leq B\right\}} \\
&= \left(S(0)e^{\sigma \widehat{W}(T)} - K\right) \mathbb{I}_{\{\widehat{W}(T) \geq k, \widehat{M}(T) \leq b\}},
\end{aligned}$$

where

$$k = \frac{1}{\sigma}\log\frac{K}{S(0)}, \quad b = \frac{1}{\sigma}\log\frac{B}{S(0)}.$$

If $k \geq 0$, we must integrate over the region $\{(m,w); k \leq w \leq m \leq b\}$. On the other hand, if $k < 0$, we integrate over the region $\{(m,w); k \leq w \leq m, 0 \leq m \leq b\}$. In both cases, the region can be described as $\{(m,w); k \leq w \leq b, w^+ \leq m \leq b\}$; see Figure 7.3.1. We assume here that $S(0) \leq B$ so that $b > 0$. Otherwise, the region over which we integrate has zero area, and the time-zero value of the call is zero rather than the integral computed below. We also assume $S(0) > 0$ so that b and k are finite. When $0 < S(0) \leq B$, the time-zero value of the up-and-out call is

$$\begin{aligned}
V(0) &= \int_k^b \int_{w^+}^b e^{-rT}\left(S(0)e^{\sigma w} - K\right) \frac{2(2m-w)}{T\sqrt{2\pi T}} e^{\alpha w - \frac{1}{2}\alpha^2 T - \frac{1}{2T}(2m-w)^2} dm dw \\
&= -\int_k^b e^{-rT}\left(S(0)e^{\sigma w} - K\right) \frac{1}{\sqrt{2\pi T}} e^{\alpha w - \frac{1}{2}\alpha^2 T - \frac{1}{2T}(2m-w)^2}\Bigg|_{m=w^+}^{m=b} dw \\
&= \frac{1}{\sqrt{2\pi T}} \int_k^b \left(S(0)e^{\sigma w} - K\right) e^{-rT + \alpha w - \frac{1}{2}\alpha^2 T - \frac{1}{2T}w^2} dw \\
&\quad - \frac{1}{\sqrt{2\pi T}} \int_k^b \left(S(0)e^{\sigma w} - K\right) e^{-rT + \alpha w - \frac{1}{2}\alpha^2 T - \frac{1}{2T}(2b-w)^2} dw \\
&= S(0)I_1 - KI_2 - S(0)I_3 + KI_4
\end{aligned}$$

where

$$\begin{aligned}
I_1 &= \frac{1}{\sqrt{2\pi T}} \int_k^b e^{\sigma w - rT + \alpha w - \frac{1}{2}\alpha^2 T - \frac{1}{2T}w^2} dw, \\
I_2 &= \frac{1}{\sqrt{2\pi T}} \int_k^b e^{-rT + \alpha w - \frac{1}{2}\alpha^2 T - \frac{1}{2T}w^2} dw, \\
I_3 &= \frac{1}{\sqrt{2\pi T}} \int_k^b e^{\sigma w - rT + \alpha w - \frac{1}{2}\alpha^2 T - \frac{1}{2T}(2b-w)^2} dw, \\
&= \frac{1}{\sqrt{2\pi T}} \int_k^b e^{\sigma w - rT + \alpha w - \frac{1}{2}\alpha^2 T - \frac{2}{T}b^2 + \frac{2}{T}bw - \frac{1}{2T}w^2} dw, \\
I_4 &= \frac{1}{\sqrt{2\pi T}} \int_k^b e^{-rT + \alpha w - \frac{1}{2}\alpha^2 T - \frac{1}{2T}(2b-w)^2} dw \\
&= \frac{1}{\sqrt{2\pi T}} \int_k^b e^{-rT + \alpha w - \frac{1}{2}\alpha^2 T - \frac{2}{T}b^2 + \frac{2}{T}bw - \frac{1}{2T}w^2} dw.
\end{aligned}$$

Remark. Similar to the vanilla options, the barrier options are still time-homogeneous under the Black-Sholes model: when we are trying to compute the option price at time t, it is equivalent to standing at time t and setting the maturity to $T - t$.

To derive the PDE for the up-and-out call option price, we need to 'truncate' the price using a stopping time: The random variable $\rho = \inf\{u \geq 0 : S(u) \geq B\}$ is a stopping time because it chooses its value based on the path of the asset price up to time ρ. The Optional Sampling Theorem, asserts that a martingale stopped at a stopping time is still a martingale. In particular, the process

$$e^{-r(t\wedge\rho)}V(t\wedge\rho)=\begin{cases}e^{-rt}V(t) & \text{if } 0\le t\le\rho\\ e^{-r\rho}V(\rho) & \text{if } \rho<t\le T\end{cases}$$

is a $\widetilde{\mathbb{P}}$-martingale.

By above analysis, we can propose that

Theorem 3.3. Let $v(t,x)$ denote the price at time t of the up-and-out call under the assumption that the call has not knocked out prior to time t and $S(t)=x$. Then $v(t,x)$ satisfies the Black-Scholes-Merton partial differential equation

$$v_t(t,x)+rxv_x(t,x)+\frac{1}{2}\sigma^2x^2v_{xx}(t,x)=rv(t,x)$$

in the rectangle $\{(t,x);0\le t<T,0\le x\le B\}$ and satisfies the boundary conditions

$$\begin{aligned}v(t,0)&=0,\quad 0\le t\le T\\ v(t,B)&=0,\quad 0\le t<T\\ v(T,x)&=(x-K)^+,\quad 0\le x\le B\end{aligned}$$

Proof. We compute the differential

$$\begin{aligned}&d\left(e^{-rt}v(t,S(t))\right)=e^{-rt}\Big[-rv(t,S(t))dt+v_t(t,S(t))dt+v_x(t,S(t))dS(t)\\ &+\frac{1}{2}v_{xx}(t,S(t))dS(t)dS(t)\Big]\\ &=e^{-rt}\left[-rv(t,S(t))+v_t(t,S(t))+rS(t)v_x(t,S(t))+\frac{1}{2}\sigma^2S^2(t)v_{xx}(t,S(t))\right]dt\\ &+e^{-rt}\sigma S(t)v_x(t,S(t))d\widetilde{W}(t)\end{aligned}$$

The dt term must be zero for $0\le t\le\rho$, (i.e., before the option knocks out). But since $(t,S(t))$ can reach any point in $\{(t,x);0\le t<T,0\le x\le B\}$ before the option knocks out, the formula must hold for every $t\in[0,T)$ and $x\in[0,B]$. □

Remark. By theorem above, we can hedge such an option theoretically. As we have

$$d\left(e^{-rt}v(t,S(t))\right)=e^{-rt}\sigma S(t)v_x(t,S(t))d\widetilde{W}(t),\quad 0\le t\le\rho.$$

The discounted value of a portfolio that at each time t holds $\Delta(t)$ shares of the underlying asset is given by

$$d\left(e^{-rt}X(t)\right)=e^{-rt}\sigma S(t)\Delta(t)d\widetilde{W}(t).$$

At least theoretically, if an agent begins with a short position in the up-and-out call and with initial capital $X(0)=v(0,S(0))$, then the usual delta-hedging formula

$$\Delta(t)=v_x(t,S(t))$$

will cause her portfolio value $X(t)$ to track the option value $v(t,S(t))$ up to the time ρ of knock-out or up to expiration T, whichever comes first.

3.2 Look-back options

An option whose payoff is based on the maximum or the minimum that the underlying asset price attains over some interval of time prior to expiration is called a lookback option. If we specify an option strike K, then this option is called a fixed-strike option, and if the strike is taken to be the price of the underlying asset at the maturity, then it is called a floating-strike option. In this section we would like to price a floating strike call option with payoff

$$\max(Y_T-S_T,0)$$

where Y_t is the running maximum of the underlying asset price. Let $t\in[0,T]$ be given. At time t, the risk-neutral price of the lookback option is

$$V(t)=\widetilde{\mathbb{E}}\left[e^{-r(T-t)}(Y(T)-S(T))\mid\mathcal{F}(t)\right]$$

Because the pair of processes $(S(t),Y(t))$ has the Markov property , there must exist a function $v(t,x,y)$ such that

$$V(t)=v(t,S(t),Y(t))$$

and this function is determined by the following boundary value problem:

Theorem 3.4. Let $v(t,x,y)$ denote the price at time t of the floating strike lookback option under the assumption that $S(t)=x$ and $Y(t)=y$. Then $v(t,x,y)$ satisfies the Black-Scholes-Merton partial differential equation

$$v_t(t,x,y)+rxv_x(t,x,y)+\frac{1}{2}\sigma^2x^2v_{xx}(t,x,y)=rv(t,x,y)$$

in the region $\{(t,x,y);0\leq t<T,0\leq x\leq y\}$ and satisfies the boundary conditions

$$\begin{aligned} v(t,0,y)&=e^{-r(T-t)}y, \quad 0\leq t\leq T, y\geq 0\\ v_y(t,y,y)&=0, \quad 0\leq t\leq T, y>0\\ v(T,x,y)&=y-x, \quad 0\leq x\leq y \end{aligned}$$

Proof. By computing the differential we have

$$\begin{aligned} &d\left(e^{-rt}v(t,S(t),Y(t))\right)\\ =&e^{-rt}\Big[-rv(t,S(t),Y(t))dt+v_t(t,S(t),Y(t))dt+v_x(t,S(t),Y(t))dS(t)+\frac{1}{2}v_{xx}(t,S(t),Y(t))dS(t)dS(t)\\ &+v_y(t,S(t),Y(t))dY(t)]\\ =&e^{-rt}\left[-rv(t,S(t),Y(t))+v_t(t,S(t),Y(t))+rS(t)v_x(t,S(t),Y(t))+\frac{1}{2}\sigma^2S^2(t)v_{xx}(t,S(t),Y(t))\right]dt\\ &+e^{-rt}\sigma S(t)v_x(t,S(t),Y(t))d\widetilde{W}(t)+e^{-rt}v_y(t,S(t),Y(t))dY(t) \end{aligned}$$

For this to be a martingale, the drift term must be zero and this gives the PDE in the boundary value problem. Also, the term $dY(t)$ is different from both $dW(t)$ and dt. This is because $Y(t)$ is continuous and nondecreasing in t. Let $0=t_0<t_1<\cdots<t_m=T$ be a partition of $[0,T]$. Then

$$\begin{aligned} &\sum_{j=1}^{m}\left(Y\left(t_j\right)-Y\left(t_{j-1}\right)\right)^2\\ &\leq\max_{j=1,\ldots,m}\left(Y\left(t_j\right)-Y\left(t_{j-1}\right)\right)\sum_{j=1}^{m}\left(Y\left(t_j\right)-Y\left(t_{j-1}\right)\right)\\ &=\max_{j=1,\ldots,m}\left(Y\left(t_j\right)-Y\left(t_{j-1}\right)\right)\cdot(Y(T)-Y(0)), \end{aligned}$$

and $\max_{j=1,\ldots,m}\left(Y\left(t_j\right)-Y\left(t_{j-1}\right)\right)$ has limit zero as $\max_{j=1,\ldots,m}\left(t_j-t_{j-1}\right)$ goes to zero because $Y(t)$ is continuous. We conclude that $Y(t)$ accumulates zero quadratic variation on $[0,T]$, a fact we record by writing

$$dY(t)dY(t)=0$$

Similarly, $dY(t)dS(t)=0$ and notice that $dY(t)$ is not a dt term: there is no process $\Theta(t)$ such that $dY(t)=\Theta(t)dt$. In other words, we cannot write $Y(t)$ as

$$Y(t)=Y(0)+\int_0^t\Theta(u)du$$

If we could, then $\Theta(u)$ would be zero whenever u is in a "flat spot" of $Y(t)$, which occurs whenever $S(t)$ drops below its maximum to date, there are time intervals in which $Y(t)$ is strictly increasing, but in fact no such interval exists. Such an interval can occur only if $S(t)$ is strictly increasing on the interval, and if there were such an interval, then $S(t)$ would accumulate zero quadratic variation on the interval. This is not the case because $dS(t)dS(t)=\sigma S^2(t)dt$ is positive for all t. Therefore, if the above representation were to hold, we would need to have $\Theta(u)=0$ for Lebesgue almost every u in $[0,T]$. This would result in $Y(t)=Y(0)$ for $0\leq t\leq T$. But in fact $Y(t)>Y(0)$ for all $t>0$, a contradiction: $dY(t)$ is not a dt term. The $dY(t)$ term is naturally zero on the "flat spots" of $Y(t)$ (i.e., when $S(t)<Y(t)$). However, at the times when $Y(t)$ increases, which are the times when $S(t)=Y(t)$, the term $e^{-rt}v_y(t,S(t),Y(t))$ must be zero because $dY(t)$ is "positive." This gives us the second boundary condition. The third boundary condition is the payoff of the option. If at any time t we have $S(t)=0$, then we will have $S(T)=0$. Furthermore, Y will be constant on $[t,T]$; if $Y(t)=y$, then $Y(T)=y$ and the price of the option at time t is this value discounted from T back to t. This gives us the first boundary condition. □

Remark. The proof of the above theorem shows that

$$d\left(e^{-rt}v(t,S(t),Y(t))\right)=e^{-rt}\sigma S(t)v_x(t,S(t),Y(t))d\widetilde{W}(t).$$

Same as before, this equation implies that the delta-hedging formula works. In contrast to the situation in the barrier option, here the function $v(t,x,y)$ is continuous and we have no problems with large delta and gamma values.

3.3 Asian options

An Asian option is one whose payoff includes a time average of the underlying asset price. The average may be over the entire time period between initiation and expiration or may be over some period of time that begins later than the initiation of the option and ends with the option's expiration. The average may be from continuous sampling,

$$\frac{1}{T}\int_0^T S(t)dt,$$

or may be from discrete sampling,

$$\frac{1}{m}\sum_{j=1}^{m} S\left(t_j\right),$$

where $0<t_1<t_2\cdots<t_m=T$.

Once again, we begin with a geometric Brownian motion $S(t)$ given by

$$dS(t)=rS(t)dt+\sigma S(t)d\widetilde{W}(t)$$

where $\widetilde{W}(t), 0\leq t\leq T$, is a Brownian motion under the risk-neutral measure $\widetilde{\mathbb{P}}$. Consider a fixed-strike Asian call whose payoff at time T is

$$V(T)=\left(\frac{1}{T}\int_0^T S(t)dt-K\right)^+$$

where the strike price K is a nonnegative constant. The price at times t prior to the expiration time T of this call is given by the risk-neutral pricing formula

$$V(t)=\widetilde{\mathbb{E}}\left[e^{-r(T-t)}V(T)\mid\mathcal{F}(t)\right],\quad 0\leq t\leq T.$$

The usual iterated conditioning argument shows that

$$e^{-rt}V(t)=\widetilde{\mathbb{E}}\left[e^{-rT}V(T)\mid\mathcal{F}(t)\right],\quad 0\leq t\leq T,$$

is a martingale under $\tilde{\mathbb{P}}$.

The Asian option payoff $V(T)$ is path-dependent. The price of the option at time t depends not only on t and $S(t)$, but also on the path that the asset price has followed up to time t. In particular, we cannot invoke the Markov property to claim that $V(t)$ is a function of t and $S(t)$ because $V(T)$ is not a function of T and $S(T); V(T)$ depends on the whole path of S. To overcome this difficulty, we augment the state $S(t)$ by defining a second process

$$Y(t)=\int_0^t S(u)du$$

The stochastic differential equation for $Y(t)$ is thus

$$dY(t)=S(t)dt$$

Because the pair of processes $(S(t),Y(t))$ is governed by the pair of stochastic differential equations, they constitute a two-dimensional Markov process. Furthermore, the call payoff $V(T)$ is a function of T and the final value $(S(T),Y(T))$ of this process. Indeed, $V(T)$ depends only on T and $Y(T)$, by the formula

$$V(T)=\left(\frac{1}{T}Y(T)-K\right)^+.$$

This implies that there must exist some function $v(t,x,y)$ such that the Asian call price is given as

$$\begin{aligned}v(t,S(t),Y(t))&=\widetilde{\mathbb{E}}\left[e^{-r(T-t)}\left(\frac{1}{T}Y(T)-K\right)^+\mid\mathcal{F}(t)\right]\\&=\widetilde{\mathbb{E}}\left[e^{-r(T-t)}V(T)\mid\mathcal{F}(t)\right].\end{aligned}$$

Theorem 3.5. The Asian call price function $v(t,x,y)$ of (7.5.7) satisfies the partial differential equation

$$\begin{gathered}v_t(t,x,y)+rxv_x(t,x,y)+xv_y(t,x,y)+\frac{1}{2}\sigma^2x^2v_{xx}(t,x,y)=rv(t,x,y)\\0\leq t<T, x\geq 0, y\in\mathbb{R}\end{gathered}$$

and the boundary conditions

$$
\begin{aligned}
v(t,0,y) &= e^{-r(T-t)}\left(\frac{y}{T}-K\right)^{+}, 0 \le t < T, y \in \mathbb{R}, \\
\lim_{y \downarrow -\infty} v(t,x,y) &= 0, 0 \le t < T, x \ge 0, \\
v(T,x,y) &= \left(\frac{y}{T}-K\right)^{+}, x \ge 0, y \in \mathbb{R}.
\end{aligned}
$$

Proof. Note that $dS(t)dY(t) = dY(t)dY(t) = 0$, we take the differential of the $\widetilde{\mathbb{P}}$-martingale $e^{-rt}V(t) = e^{-rt}v(t,S(t),Y(t))$. This differential is

$$
\begin{aligned}
& d\left(e^{-rt}v(t,S(t),Y(t))\right) \\
&= e^{-rt}\left[-rvdt + v_t dt + v_x dS + v_y dY + \frac{1}{2}v_{xx}dSdS\right] \\
&= e^{-rt}\left[-rv + v_t + rSv_x + Sv_y + \frac{1}{2}\sigma^2 S^2 v_{xx}\right]dt + e^{-rt}\sigma S v_x d\widetilde{W}(t).
\end{aligned}
$$

In order for this to be a martingale, the dt term must be zero, which implies

$$
\begin{aligned}
v_t(t,S(t),Y(t)) + rS(t)v_x(t,S(t),Y(t)) + S(t)v_y(t,S(t),Y(t)) \\
+\frac{1}{2}\sigma^2 S^2(t) v_{xx}(t,S(t),Y(t)) = rv(t,S(t),Y(t))
\end{aligned}
$$

Replacing $S(t)$ by the dummy variable x and $Y(t)$ by the dummy variable y, we obtain the pde.
We note that $S(t)$ must always be nonnegative, and so the pde holds for $x \ge 0$. If $S(t) = 0$ and $Y(t) = y$ for some value of t, then $S(u) = 0$ for all $u \in [t,T]$, and so $Y(u)$ is constant on $[t,T]$. Therefore, $Y(T) = y$, and the value of the Asian call at time t is $\left(\frac{y}{T}-K\right)^{+}$, discounted from T back to t. This gives us the first boundary condition.
In contrast, it is not the case that if $Y(t) = 0$ for some time t, then $Y(u) = 0$ for all $u \ge 0$. Therefore, we cannot easily determine the value of $v(t,x,0)$, and we do not provide a condition on the boundary $y = 0$. Indeed, at least mathematically there is no problem with allowing y to be negative. If at time t we set $Y(t) = y$, then

$$
Y(T) = y + \int_t^T S(u)du
$$

Even if y is negative, this makes sense, and in this case we could still have $Y(T) > 0$ or even $\frac{1}{T}Y(T) - K > 0$, so that the call expires in the money. When using the differential equations to describe the "state" processes $S(t)$ and $Y(t)$, there is no reason to require that $Y(t)$ be nonnegative. (We still require that $S(t)$ be nonnegative because $x = 0$ is a natural boundary for $S(t)$.) For this reason, we do not restrict the values of y in the partial differential equation. The natural boundary for y is $y = -\infty$. If $Y(t) = y, S(t) = x$, and holding x fixed we let $y \to -\infty$, then $Y(T)$ approaches $-\infty$, the probability that the call expires in the money approaches zero, and the option price approaches zero. The natural boundary for y is $y = -\infty$, and the boundary condition there is the second boundary condition. The third boundary condition is just the payoff of the call. □

Remark. The Delta-hedging still works: After we set the dt term equal to zero, we see that

$$
d\left(e^{-rt}v(t,S(t),Y(t))\right) = e^{-rt}\sigma S(t)v_x(t,S(t),Y(t))d\widetilde{W}(t)
$$

The discounted value of a portfolio that at each time t holds $\Delta(t)$ shares of the underlying asset is given by

$$
d\left(e^{-rt}X(t)\right) = e^{-rt}\sigma S(t)\Delta(t)d\widetilde{W}(t)
$$

To hedge a short position in the Asian call, an agent should equate these two differentials, which leads to the delta-hedging formula

$$
\Delta(t) = v_x(t,S(t),Y(t))
$$

3.4 American options: perceptual options

The simplest interesting American option is the perpetual American put. It is interesting because the optimal exercise policy is not obvious, and it is simple because this policy can be determined explicitly. Although this is not a traded option, we begin our discussion with it in order to present in a simple context the ideas behind the subsequent analysis of more realistic options.

The underlying asset in most of this chapter has the price process $S(t)$ given by

$$
dS(t) = rS(t)dt + \sigma S(t)d\widetilde{W}(t)
$$

where the interest rate r and the volatility σ are strictly positive constants and $\widetilde{W}(t)$ is a Brownian motion under the risk-neutral probability measure $\widetilde{\mathbb{P}}$. The perpetual American put pays $K - S(t)$ if it is exercised at time t. This is its intrinsic value.

Definition 3.6. Let $\mathcal{T}$ be the set of all stopping times. The price of the perpetual American put is defined to be

$$v_*(x) = \max_{\tau \in \mathcal{T}} \widetilde{\mathbb{E}}\left[e^{-r\tau}(K - S(\tau))\right],$$

where $x = S(0)$is the initial stock price. In the event that $\tau = \infty$, we interpret $e^{-r\tau}(K - S(\tau))$ to be zero.

Theorem 3.7. (Laplace transform for first passage time of drifted Brownian motion). Let $\widetilde{W}(t)$ be a Brownian motion under a probability measure $\tilde{\mathbb{P}}$, let μ be a real number, and let m be a positive number. Define $X(t) = \mu t + \widetilde{W}(t)$, and set

$$\tau_m = \min\{t \geq 0; X(t) = m\},$$

so that τ_m is the stopping time of Example 8.2.3. If $X(t)$ never reaches the level m, then we interpret τ_m to be ∞. Then

$$\tilde{\mathbf{E}} e^{-\lambda \tau_m} = e^{-m\left(-\mu + \sqrt{\mu^2 + 2\lambda}\right)} \text{ for all } \lambda > 0,$$

where we interpret $e^{-\lambda\tau_m}$ to be zero if $\tau_m = \infty$.

This is a problem in the midterm, so we omit its proof here.

Suppose the owner of the perpetual American put sets a positive level $L < K$ and resolves to exercise the put the first time the stock price falls to L. If the initial stock price is at or below L, she exercises immediately (at time zero). The value of the put in this case is $v_L(S(0)) = K - S(0)$. If the initial stock price is above L, she exercises at the stopping time

$$\tau_L = \min\{t \geq 0; S(t) = L\}$$

where τ_L is set equal to ∞ if the stock price never reaches the level L. At the time of exercise, the put pays $K - S(\tau_L) = K - L$. Discounting this back to time zero and taking the risk-neutral expected value, we compute the value of the put under this exercise strategy to be

$$v_L(S(0)) = (K - L)\widetilde{\mathbb{E}} e^{-r_\tau} \text{ for all } S(0) \geq L$$

On those paths where $\tau_L = \infty$, we interpret $e^{-r\tau_L}$ to be zero. (Recall our assumption at the beginning of this section that r is strictly positive.) Although not explicitly indicated by the notation, the distribution of τ_L depends on the initial stock price $S(0)$, so the right-hand side (8.3.10) is a function of $S(0)$.

Theorem 3.8. The function $v_L(x)$ is given by the formula

$$v_L(x) = \begin{cases} K - x, & 0 \leq x \leq L \\ (K - L)\left(\frac{x}{L}\right)^{-\frac{2r}{\sigma^2}}, & x \geq L \end{cases}$$

Proof. We consider the case $S(0) = x > L$. The stopping time τ_L is the first time

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

reaches the level L. But $S(t) = L$ if and only if

$$-\widetilde{W}(t) - \frac{1}{\sigma}\left(r - \frac{1}{2}\sigma^2\right)t = \frac{1}{\sigma}\log\frac{x}{L}.$$

We now apply the theorem above with $X(t)$ in that theorem replaced by $-\widetilde{W}(t) - \frac{1}{\sigma}\left(r - \frac{1}{2}\sigma^2\right)t$ (the processes $\widetilde{W}(t)$ and $-\widetilde{W}(t)$ are both Brownian motions under $\widetilde{\mathbb{P}}$), with λ replaced by r, with μ replaced by $-\frac{1}{\sigma}\left(r - \frac{1}{2}\sigma^2\right)$, and with m replaced by $\frac{1}{\sigma}\log\frac{x}{L}$, which is positive. With these replacements, τ_m in the theorem is τ_L and

$$\begin{aligned}
\mu^2 + 2\lambda &= \frac{1}{\sigma^2}\left(r^2 - r\sigma^2 + \frac{1}{4}\sigma^4\right) + 2r \\
&= \frac{1}{\sigma^2}\left(r^2 + r\sigma^2 + \frac{1}{4}\sigma^4\right) \\
&= \frac{1}{\sigma^2}\left(r + \frac{1}{2}\sigma^2\right)^2,
\end{aligned}$$

Therefore,

$$-\mu + \sqrt{\mu^2 + 2\lambda} = \frac{1}{\sigma}\left(r - \frac{1}{2}\sigma^2\right) + \frac{1}{\sigma}\left(r + \frac{1}{2}\sigma^2\right) = \frac{2r}{\sigma}.$$

and hence

$$\widetilde{\mathbb{E}}^{-r\tau_L} = \exp\left\{-\frac{1}{\sigma}\log\frac{x}{L}\cdot\frac{2r}{\sigma}\right\} = \left(\frac{x}{L}\right)^{-\frac{2r}{\sigma^2}}.$$

□

We must first determine the value of L_*. We note that

$$v_L(x) = (K-L)L^{\frac{2r}{\sigma^2}}x^{-\frac{2r}{\sigma^2}} \text{ for all } x \geq L$$

We know that L_* is the value of L that maximizes this quantity when we hold x fixed. We thus define

$$g(L) = (K-L)L^{\frac{2r}{\sigma^2}}$$

and seek the value of L that maximizes this function over $L \geq 0$. Because $\frac{2r}{\sigma^2}$ is strictly positive, we have $g(0)=0$ and $\lim_{L\to\infty} g(L) = -\infty$. Moreover,

$$g'(L) = -L^{\frac{2r}{\sigma^2}} + \frac{2r}{\sigma^2}(K-L)L^{\frac{2r}{\sigma^2}-1} = -\frac{2r+\sigma^2}{\sigma^2}L^{\frac{2r}{\sigma^2}} + \frac{2r}{\sigma^2}KL^{\frac{2r}{\sigma^2}-1}.$$

Setting this equal to zero, we solve for

$$L_* = \frac{2r}{2r+\sigma^2}K.$$

This is a number between 0 and K. Furthermore,

$$g(L_*) = \frac{\sigma^2}{2r+\sigma^2}\left(\frac{2r}{2r+\sigma^2}\right)^{\frac{2r}{\sigma^2}} K^{\frac{2r+\sigma^2}{\sigma^2}}$$

We have

$$v_{L*}(x) = \begin{cases} K-x, & 0 \leq x \leq L_*, \\ (K-L_*)\left(\frac{x}{L_*}\right)^{-\frac{2r}{\sigma^2}}, & x \geq L_*, \end{cases}$$

so that

$$v'_{L_*}(x) = \begin{cases} -1, & 0 \leq x \leq L_*, \\ -(K-L_*)\frac{2r}{\sigma^2 x}\left(\frac{x}{L_*}\right)^{-\frac{2r}{\sigma^2}}, & x \geq L_*. \end{cases}$$

The second derivative of $v(x)$ has a jump at $x = L_*$, and hence is undefined at this point. Indeed,

$$v''_{L_*}(x) = \begin{cases} 0, & 0 \leq x < L_*, \\ (K-L_*)\frac{2r\left(2r+\sigma^2\right)}{\sigma^4 x^2}\left(\frac{x}{L_*}\right)^{-\frac{2r}{\sigma^2}}, & x > L_*. \end{cases}$$

The left-hand and right-hand second derivatives at $x = L_*$ are $v\left(L_*-\right) = 0$ and $v''\left(L_*+\right) = (K-L_*)\frac{2r\left(2r+\sigma^2\right)}{\sigma^4 L^2} > 0$. For $x > L_*$, we can verify by direct computation that

$$\begin{aligned} & rv_{L_*}(x) - rxv'_{L_*}(x) - \frac{1}{2}\sigma^2 x^2 v''_{L_*}(x) \\ & = (K-L_*)\left(r + \frac{2r^2}{\sigma^2} - \frac{r\left(2r+\sigma^2\right)}{\sigma^2}\right)\left(\frac{x}{L_*}\right)^{-\frac{2r}{\sigma^2}} = 0 \end{aligned}$$

On the other hand, for $0 \leq x < L_*$,

$$rv_{L_*}(x) - rxv'_{L.}(x) - \frac{1}{2}\sigma^2 x^2 v''_{L.}(x) = r(K-x) + rx = rK.$$

In particular, we see that $v_{L_*}(x)$ satisfies the so-called linear complementarity conditions

$$\begin{aligned} v(x) &\geq (K-x)^+ \text{for all } x \geq 0, \\ rv(x) - rxv'(x) - \frac{1}{2}\sigma^2 x^2 v''(x) &\geq 0 \text{ for all } x \geq 0, \text{ and} \end{aligned}$$

for each $x \geq 0$, equality holds in either case. The point L_* is slightly problematical in (8.3.19) since $v''_{L.}(L_*)$ is undefined. However, if we replace $v''_{L_*}(L_*)$ in (8.3.19) by either $v''_{L_*}\left(\dot{L}_*-\right)$ or $v''_{L_*}(L_*+)$, the inequality holds.

Theorem 3.9. Let $S(t)$ be the stock price given by (8.3.1) and let $\tau_{L.}$ be given by (8.3.9) with $L = L_*$. Then $e^{-rt^-}v_{L_*}(S(t))$ is a supermartingale under $\widetilde{\mathbb{P}}$, and the stopped process $e^{-r(t\wedge\tau_{L_*})_{L_*}}\left(S\left(t\wedge\tau_{L_*}\right)\right)$ is a martingale.

Proof. Fortunately, the Itô-Doeblin formula applies to functions whose second derivatives have jumps, provided the first derivative is continuous. We may thus compute

$$\begin{aligned}
& d\left[e^{-rt} v_{L_*}(S(t))\right] \\
& = e^{-rt}\left[-r v_{L_*}(S(t)) dt + v_{L_*}^{\prime}(S(t)) dS(t) + \frac{1}{2} v_{L_*}^{\prime\prime}(S(t)) dS(t) dS(t)\right] \\
& = e^{-rt}\left[-r v_{L_*}(S(t)) + r S(t) v_{L_*}^{\prime}(S(t)) + \frac{1}{2} \sigma^2 S^2(t) v_{L_*}^{\prime\prime}(S(t))\right] dt \\
& + e^{-rt} \sigma S(t) v_{L_*}^{\prime}(S(t)) d\widetilde{W}(t)
\end{aligned}$$

Because of the pde, the dt term in this expression is either 0 or $-rK$, depending on whether $S(t) > L_*$ or $S(t) < L_*$. If $S(t) = L_*, v_{L_*^{\prime\prime}}(S(t))$ is undefined, but the probability $S(t) = L_*$ is zero so this does not matter. We thus have

$$d\left[e^{-rt} v_{L_*}(S(t))\right] = -e^{-rt} r K \mathbb{H}_{\{S(t)<L^*\}} dt + e^{-rt} \sigma S(t) v_{L_*}^{\prime}(S(t)) d\widetilde{W}(t)$$

(8.3.21) Because the dt term is less than or equal to zero, $e^{-rt} v_{L_*}(S(t))$ is a supermartingale; when $S(t) < L_*$ it has a downward tendency. If the initial stock price is above L_*, then prior to the time τ_{L_*} when the stock price first reaches L_*, the dt term in (8.3.21) is zero and hence $e^{-r(t\wedge\tau_{L_*})} v\left(S\left(t\wedge\tau_{L_*}\right)\right)$ is a martingale. Indeed, integration yields

$$e^{-r(t\wedge\tau_{L_*})} v_{L_*}\left(S\left(t\wedge\tau_{L_*}\right)\right) = v_{L_*}(0) + \int_0^{t\wedge\tau_{L_*}} e^{-ru} \sigma S(u) v_{L_*}^{\prime}(S(u)) d\widetilde{W}(u)$$

Itô integrals are martingales, and hence the Itô integral above stopped at the stopping time τ_L, is a martingale. ☐

Corollary 3.10. Recall that $\mathcal{T}$ is the set of all stopping times, not just those of the form (8.3.9). We have

$$v_{L^\cdot}(x) = \max_{\tau\in T} \widetilde{\mathbb{E}}\left[e^{-r\tau}(K - S(\tau))\right]$$

where $x = S(0)$ is the initial stock price. In other words, $v_{L_*}(x)$ is the perpetual American put price.

Proof. Because $e^{-rt} v_{L_*}(S(t))$ is a supermartingale under $\tilde{\mathbb{P}}$, we have from optional sampling theorem that, for every stopping time $\tau \in \mathcal{T}$,

$$v_{L_*}(x) = v_{L_*}(S(0)) \geq \widetilde{\mathbb{E}}\left[e^{-r(t\wedge\tau)} v_{L_*}(S(t\wedge\tau))\right].$$

Because $v_{L^\cdot}(S(t\wedge\tau))$ is bounded, we may let $t \to \infty$, using the Dominated Convergence Theorem to conclude that

$$v_{L^\cdot}(x) \geq \widetilde{\mathbb{E}}\left[e^{-r\tau} v_{L^\cdot}(S(\tau))\right] \geq \tilde{\mathbb{E}}\left[e^{-r\tau}(K - S(\tau))\right],$$

. Because this inequality holds for every $\tau \in \mathcal{T}$, we have

$$v_{L^\cdot}(x) \geq \max_{\tau\in\mathcal{T}} \widetilde{\mathbb{E}}\left[e^{-r\tau}(K - S(\tau))\right].$$

On the other hand, if we replace τ by $\tau_{L^\cdot}$, we obtain equality because $e^{-r(t\wedge\tau_{L_*})} v\left(S\left(t\wedge\tau_{L_*}\right)\right)$ is a martingale under $\widetilde{\mathbb{P}}$. Letting $t \to \infty$ and using the Dominated Convergence Theorem, we obtain

$$v_{L^\cdot}(x) = \tilde{\mathbb{E}}\left[e^{-r_{L^\cdot}} v_{L^\cdot}\left(S\left(\tau_{L^\cdot}\right)\right)\right]$$

Since

$$e^{-r\tau_{L^\cdot}} v_{L_*}\left(S\left(\tau_{L_*}\right)\right) = e^{-r\tau_L} v_{L_*}\left(L_*\right) = e^{-r\tau_L}\left(K - L_*\right) = e^{-r\tau_L}\left(K - S\left(\tau_{L_*}\right)\right)$$

if $\tau_{L_*} < \infty$ (and is interpreted to be zero if $\tau_{L^\cdot} = \infty$), we see that

$$v_{L*}(x) = \widetilde{\mathbb{E}}\left[e^{-r\tau_L}\left(K - S\left(\tau_{L_*}\right)\right)\right].$$

It follows that

$$v_{L^\cdot}(x) \leq \max_{\tau\in T} \widetilde{\mathbb{E}}\left[e^{-r\tau}(K - S(\tau))\right]$$

☐

Remark. Discounted European option prices are martingales under the risk-neutral probability measure. Discounted American option prices are martingales up to the time they should be exercised. If they are not exercised when they should be, they tend downward. Since a martingale is a special case of a supermartingale, and processes that tend downward are supermartingales, discounted American option prices are supermartingales. An agent who is short an American option can hedge that short position in the usual way during the time the discounted option price is a martingale. If the option is not exercised when it should be, then the agent can continue the hedge and take money off the table. The following corollary illustrates this for the perpetual American put of this section.

Corollary 3.11. Consider an agent with initial capital $X(0) = v_{L.}(S(0))$, the initial perpetual American put price. Suppose this agent uses the portfolio process $\Delta(t) = v'_{L_*}(S(t))$ and consumes cash at rate $C(t) = rK\mathbb{I}_{\{S(t)<L^*\}}$ (i.e., consumes cash at rate rK whenever $S(t) < L^*$). Then the value $X(t)$ of the agent's portfolio agrees with the option price $v_{L.}(S(t))$ for all times t until the option is exercised. In particular, $X(t) \geq (K - S(t))^+$for all t until the option is exercised, so the agent can pay off a short option position regardless of when the option is exercised.

Proof. The differential of the agent's portfolio value process is

$$dX(t) = \Delta(t)dS(t) + r(X(t) - \Delta(t)S(t))dt - C(t)dt$$

so the differential of the discounted portfolio value process is

$$\begin{aligned} d\left(e^{-rt}X(t)\right) &= e^{-rt}(-rX(t)dt + dX(t)) \\ &= e^{-rt}(\Delta(t)dS(t) - r\Delta(t)S(t)dt - C(t)dt) \\ &= e^{-rt}(\Delta(t)\sigma S(t)d\widetilde{W}(t) - C(t)dt). \end{aligned}$$

Substituting $\Delta(t) = v'_{L*}(S(t))$ and $C(t) = rK\mathbb{I}_{\{S(t)<L*\}}$ into (8.3.24) and comparing it to (8.3.21), we see that $d\left(e^{-rt}X(t)\right) = d\left[e^{-rt}v_L(S(t))\right]$. Integrating both sides of this equation and using the initial equality $X(0) = v_{L_*}(S(0))$, we obtain $X(t) = v_{L_*}(S(t))$ for all t prior to exercise. □

Remark. During any period in which $S(t) < L^*$, the agent in the above corollary has stock position $\Delta(t) = v'_{L.}(S(t)) = -1$ (i.e., is short one share of stock) and has a total portfolio value $X(t) = v_{L.}(S(t)) = K - S(t)$. Therefore, the agent has K invested in the money market. If the owner of the put exercises, the agent receives a share of stock, which covers his short position, and pays out K from his money market account. If the owner of the put does not exercise, the agent holds his position and consumes the interest from the money market investment (i.e., consumes cash at rate rK per unit time).

The argument in applies generally. In a complete market, whenever some discounted price process is a supermartingale, it is possible to construct a hedging portfolio whose value tracks the price process. This portfolio may sometimes consume. If, in addition, the price process dominates some intrinsic value, then a short position in the American option with that intrinsic value can be hedged. There are always two conditions on the price of any American option. These conditions guarantee that the price is sufficient to satisfy the seller of the put.

However, the inequality conditions and alone are not enough to determine the price of the perpetual American put. There can be functions that satisfy these conditions but are strictly greater than the price $v_{L*}(x)$. There must be some additional condition that guarantees that the price is satisfactory for the purchaser of the put. One version of this condition for the perpetual American put is the equality condition. It guarantees that there exists an exercise strategy that permits the owner of the put to capture the full value of the put.

3.5 American options: finite maturity

Definition 3.12. Let $0 \leq t \leq T$ and $x \geq 0$ be given. Assume $S(t) = x$. Let $\mathcal{F}_u^{(t)}, t \leq u \leq T$, denote the σ-algebra generated by the process $S(v)$ as v ranges over $[t, u]$, and let $\mathcal{T}_{t,T}$ denote the set of stopping times for the filtration $\mathcal{F}_u^{(t)}, t \leq u \leq T$, taking values in $[t, T]$ or taking the value ∞. times t and u. The price at time t of the American put expiring at time T is defined to be

$$v(t,x) = \max_{\tau \in \mathcal{T}_{t,T}} \widetilde{\mathbb{E}}\left[e^{-r(\tau - t)}(K - S(\tau)) \mid S(t) = x\right].$$

In the event that $\tau = \infty$, we interpret $e^{-r\tau}(K - S(\tau))$ to be zero. This is the case when the put expires unexercised.

The finite-expiration American put price function $v(t, x)$ satisfies the linear complementarity conditions (cf.

$$v(t,x) \geq (K - x)^+ \text{for all } t \in [0,T], x \geq 0$$

$$rv(t,x) - v_t(t,x) - rxv_x(t,x) - \frac{1}{2}\sigma^2 x^2 v_{xx}(t,x) \geq 0$$

$$\text{for all } t \in [0,T), x \geq 0, \text{ and}$$

for each $t \in [0, T)$ and $x \geq 0$, equality holds in either case.

Theorem 3.13. Let $S(u), t \leq u \leq T$, be the stock price starting at $S(t) = x$ and with the stopping set $\mathcal{S}$. Let

$$\tau_* = \min\{u \in [t,T]; (u, S(u)) \in \mathcal{S}\},$$

where we interpret τ_* to be ∞ if $(u, S(u))$ doesn't enter $\mathcal{S}$ for any $u \in [t, T]$. Then $e^{-ru}v(u, S(u)), t \leq u \leq T$, is a supermartingale under $\widetilde{\mathbb{P}}$, and the stopped process $e^{-r(u\wedge\tau_*)}v\left(u, S\left(u\wedge\tau_*\right)\right), t \le u \le T$, is a martingale.

Proof. The Itô-Doeblin formula applies to $e^{-ru}v(u, S(u))$ showing that

$$\begin{aligned}
&d\left[e^{-ru}v(u, S(u))\right] \\
&= e^{-ru}\Big[-rv(u, S(u))du + v_u(u, S(u))du + v_x(u, S(u))dS(u) \\
&+\frac{1}{2}v_{xx}(u, S(u))dS(u)dS(u)\Big] \\
&= e^{-ru}\Big[-rv(u, S(u)) + v_u(u, S(u)) + rS(u)v_x(u, S(u)) \\
&+\frac{1}{2}\sigma^2S^2(u)v_{xx}(u, S(u))\Big]du + e^{-ru}\sigma S(u)v_x(u, S(u))d\widetilde{W}(u).
\end{aligned}$$

The du term inis $-e^{-ru}rK\mathbb{I}_{\{S(u)<L(T-u)\}}$. This is nonpositive, and so $e^{-ru}v(u, S(u))$ is a supermartingale under $\widetilde{\mathbb{P}}$. In fact, starting from $u = t$ and up until time τ_*, we have $S(u) > L(T-u)$, so the du term is zero. Therefore, the stopped process $e^{-r(u\wedge\tau_*)}v\left(u\wedge\tau_*, S\left(u\wedge\tau_*\right)\right)$, $t \le u \le T$, is a martingale. □

Corollary 3.14. Consider an agent with initial capital $X(0) = v(0, S(0))$, the initial finite-expiration put price. Suppose this agent uses the portfolio process $\Delta(u) = v_x(u, S(u))$ and consumes cash at rate $C(u) = rK\mathbb{I}_{\{S(u)<L(T-u)\}}$ per unit time. Then $X(u) = v(u, S(u))$ for all times u between $u = 0$ and the time the option is exercised or expires. In particular, $S(u) \ge (K - S(u))^+$for all times u until the option is exercised or expires, so the agent can pay off a short option position regardless of when the option is exercised.

Now we consider American call options. We begin with a case slightly more general than a call option. Consider a stock whose price process $S(t)$ is given by

$$dS(t) = rS(t)dt + \sigma S(t)d\widetilde{W}(t)$$

where the interest rate r and the volatility σ are strictly positive and $\widetilde{W}(t)$ is a Brownian motion under the risk-neutral probability measure $\widetilde{\mathbb{P}}$.

Lemma 3.15. Let $h(x)$ be a nonnegative, convex function of $x \ge 0$ satisfying $h(0) = 0$. Then the discounted intrinsic value $e^{-rt}h(S(t))$ of the American derivative security that pays $h(S(t))$ upon exercise is a submartingale.

Proof. Because $h(x)$ is convex, for $0 \le \lambda \le 1$ and $0 \le x_1 \le x_2$, we have

$$h\left((1-\lambda)x_1 + \lambda x_2\right) \le (1-\lambda)h\left(x_1\right) + \lambda h\left(x_2\right).$$

Taking $x_1 = 0, x_2 = x$, and using the fact that $h(0) = 0$, we obtain that

$$h(\lambda x) \le \lambda h(x) \text{ for all } x \ge 0, 0 \le \lambda \le 1.$$

For $0 \le u \le t \le T$, we have $0 \le e^{-r(t-u)} \le 1$, and implies

$$\widetilde{\mathbb{E}}\left[e^{-r(t-u)}h(S(t)) \mid \mathcal{F}(u)\right] \ge \widetilde{\mathbb{E}}\left[h\left(e^{-r(t-u)}S(t)\right) \mid \mathcal{F}(u)\right].$$

The conditional Jensen's inequality implies

$$\begin{aligned}
\widetilde{\mathbb{E}}\left[h\left(e^{-r(t-u)}S(t)\right) \mid \mathcal{F}(u)\right] &\ge h\left(\tilde{\mathbb{E}}\left[e^{-r(t-u)}S(t) \mid \mathcal{F}(u)\right]\right) \\
&= h\left(e^{ru}\widetilde{\mathbb{E}}\left[e^{-rt}S(t) \mid \mathcal{F}(u)\right]\right).
\end{aligned}$$

Because $e^{-rt}S(t)$ is a martingale under $\tilde{\mathbb{P}}$, we have

$$h\left(e^{ru}\widetilde{\mathbb{E}}\left[e^{-rt}S(t) \mid \mathcal{F}(u)\right]\right) = h\left(e^{ru}e^{-ru}S(u)\right) = h(S(u)).$$

Putting all these together, we conclude that

$$\widetilde{\mathbb{E}}\left[e^{-r(t-u)}h(S(t)) \mid \mathcal{F}(u)\right] \ge h(S(u))$$

or, equivalently,

$$\widetilde{\mathbb{E}}\left[e^{-rt}h(S(t)) \mid \mathcal{F}(u)\right] \ge e^{-ru}h(S(u)).$$

This is the submartingale property for $e^{-rt}h(S(t))$. □

Theorem 3.16. Let $h(x)$ be a nonnegative, convex function of $x \geq 0$ satisfying $h(0) = 0$. Then the price of the American derivative security expiring at time T and having intrinsic value $h(S(t)), 0 \leq t \leq T$, is the same as the price of the European derivative security paying $h(S(T))$ at expiration T. In particular, the call options satisfy these conditions.

4 Term-structure models

4.1 Introduction

On the fixed-income market, we can only observe bond prizes instead of the actual interest rate. These bonds can be decomposed int zero-coupon bonds with different maturities, and for these zero-coupon bonds we can determine their yield-to-maturity and hence the price for forward.

The modeling procedure is as follows: let $B(t,T)$ be the zero-coupon bond price at time T and maturity T with face value 1, then the yield-to-maturity $y(t,T)$ can be computed as

$$\begin{aligned} e^{y(t,T)(T-t)} &= \frac{1}{B(t,T)} \\ \Rightarrow y(t,T) &= -\frac{\ln(B(t,T))}{T-t} \end{aligned}$$

Then for another bond paying coupons. $c_1, c_2, \ldots, c_n$ at time $T_1, T_2, \ldots, T_x$, its fair price at tile t is

$$P_n = \sum_{i=1}^{n} c_i B\left(t, T_i\right)$$

After observing these P_n, we can do bootstrapping.

$$\begin{aligned} P_1 &= C_{1,1} B\left(t, T_1\right) \\ P_2 &= C_{2,1} B\left(t, T_1\right) + C_{2,2} B\left(t, T_2\right) \\ P_n &= C_{n,1} B\left(t, T_1\right) + \cdots + C_{n,n} B\left(t, T_n\right) \end{aligned}$$

By solute this linear system we can obtain $B\left(t, T_i\right) \forall 1 \leq i \leq n$. and hence $y\left(t, T_i\right) = -\frac{\ln B\left(t_1 T_i\right)}{T_i - t}$

Then we can plot the yield curve for the function $T \rightarrow y(t,T)$, which is usually upward slopping (and hence long time bonds provide more benefits).

Now we instigate the short rate and the forward rate. Short rate is the yield-to-maturity for the shortest maturity, and forward rate is the interest rate of lending from T_1 to T_2 sturdily at time t, denoted as $f\left(t, T_1, T_2\right)$.

Suppose at time t we short $\frac{B\left(t_1 T_1\right)}{B\left(t_1 T_2\right)}$ shares of zero-coupon-bonds matured at T_2 and buy 1 share of zero-coupon-bond matured at T_1. Then at T_1 we receive 1 dollar and until T_2 we need to pay $\frac{B(t_1(T)}{B_1 T_2)}$. Hence $f\left(t_1 T_1, T_2\right)$ is given by

$$\begin{aligned} 1 + f\left(t, T_1, T_2\right)\left(T_2 - T_1\right) &= \frac{B\left(t, T_1\right)}{B\left(t, T_2\right)} \\ \Rightarrow f\left(t, T_1, T_2\right) &= \frac{B\left(t, T_1\right) - B(t, T_2))}{\left(T_2 - T_1\right)} \cdot \frac{1}{B\left(t, T_2\right)} \end{aligned}$$

Hence the interest rate of lending at time T standing at time is

$$\begin{aligned} f(t,T) &= \lim_{T_1, T_2 \rightarrow T} f\left(t, T_1 T_2\right) = -\frac{\partial}{\partial T} B(t,T) \cdot \frac{1}{B(t,T)} = -\frac{\partial}{\partial T} \ln B(t,T) \\ \Rightarrow B(t,T) &= \exp\left(-\int_t^T f(t,v) dv\right) \end{aligned}$$

and hence the short rate is given by $r_t = f(t,t)$.

After these introductions we propose two methods to model the market: the first way is calibrate the parameters from the market and then we specify the interest rate directly. In this way there are also two kinds of models: the one-factor models that directly specifies the dynamics of r_t and the two-factor models that relates r_t with 'factors' by an affine function (which are called affine-yield models): The second way is that we model the forward rate directly with out specifying r_t, which can also determine the price of the derivatives.

4.2 Affine-yield models

We have two classical one-factor models:

(1) Hull-white model:

$$dr_t = k\left(\theta - r_t\right) dt + \sigma d\widetilde{W}_t, k, \theta\sigma > 0$$

(2) CIR model:

$$dr_t = k\left(\theta - r_t\right) dt + \sigma\sqrt{r_t} da_t, : k_1\theta, \sigma > 0$$

Both models are mean-reverting, but for Hull-white model, $r_{t \in \mathbb{R}}$, which is not realistic. For CIR model, $r_t \geqslant 0$ as when r_t decreases to zero, the term $\sigma\sqrt{r_t} d\widetilde{W}_t$ becomes small and the drift term would drive it back to positive.

The zero-coupon-bond prices can be determine by pdes:

$$\begin{aligned} B(t,T) &= \widetilde{\mathbb{E}}\left[e^{-\int_t^T r_s ds} \mid \mathcal{F}_t\right] \\ &= \widetilde{\mathbb{E}}\left[e^{-\int_t^T r_s ds} \mid r_t\right] := v\left(t, r_t\right) \end{aligned}$$

By Feymann-Kac formula,

$$\begin{cases} V_t + \mathcal{A}_t v - rv = 0 \\ v(T,r) = 1 \end{cases}$$

where $\mathcal{A}_t$ is generator of the sde.

The solutions to the BVP is of the form $v(t,r) = e^{-A(t)-rC(t)}$ and hence the yield-to-maturity is given by $y(t,T) = -\frac{\ln(B,T)}{T-t} = \frac{1}{T-t}\left(A(t) + r_t c(t)\right)$, which is an affine function of r_t (which is the reason of the name affine-yield models)

For the two-factor Vasicek model, we let the factors $X_1(t)$ and $X_2(t)$ be given by the system of stochastic differential equations

$$\begin{aligned} dX_1(t) &= \left(a_1 - b_{11}X_1(t) - b_{12}X_2(t)\right)dt + \sigma_1 d\widetilde{B}_1(t), \\ dX_1(t) &= \left(a_2 - b_{21}X_1(t) - b_{22}X_2(t)\right)dt + \sigma_2 d\widetilde{B}_2(t) \\ R(t) &= \epsilon_0 + \epsilon_1 X_1(t) + \epsilon_2 X_2(t) \end{aligned}$$

where the processes $\widetilde{B}_1(t)$ and $\widetilde{B}_2(t)$ are Brownian motions under a risk-neutral measure $\widetilde{\mathbb{P}}$ with constant correlation $\nu \in (-1,1)$

Note that this model is over-parametrized, so we introduce its canonical form

$$\begin{aligned} dY_1(t) &= -\lambda_1 Y_1(t)dt + d\widetilde{W}_1(t) \\ dY_2(t) &= -\lambda_{21}Y_1(t)dt - \lambda_2 Y_2(t)dt + d\widetilde{W}_2(t) \\ R(t) &= \delta_0 + \delta_1 Y_1(t) + \delta_2 Y_2(t) \end{aligned}$$

where $\widetilde{W}_1(t)$ and $\widetilde{W}_2(t)$ are independent Brownian motions.

We have also the two-factor CIR model: The evolution of the factor processes in the canonical two-factor CIR model is given by

$$\begin{aligned} dY_1(t) &= \left(\mu_1 - \lambda_{11}Y_1(t) - \lambda_{12}Y_2(t)\right)dt + \sqrt{Y_1(t)}d\widetilde{W}_1(t) \\ dY_2(t) &= \left(\mu_2 - \lambda_{21}Y_1(t) - \lambda_{22}Y_2(t)\right)dt + \sqrt{Y_2(t)}d\widetilde{W}_2(t) \end{aligned}$$

assuming that

$$\mu_1 \geq 0, \mu_2 \geq 0, \lambda_{11} > 0, \lambda_{22} > 0, \lambda_{12} \leq 0, \lambda_{21} \leq 0$$

and the interest rate process is given by

$$R(t) = \delta_0 + \delta_1 Y_1(t) + \delta_2 Y_2(t)$$

Here the assumptions $\mu_1 \geq 0, \mu_2 \geq 0, \lambda_{12} \leq 0, \lambda_{21} \leq 0$ guarantee that the long-run interest rate is positive and the assumptions $\lambda_{12} \leq 0, \lambda_{21} \leq 0$ guarantee that the process is mean-reverting.

We also have the mixture model: The canonical two-factor mixed model is

$$\begin{aligned} dY_1(t) &= \left(\mu - \lambda_1 Y_1(t)\right)dt + \sqrt{Y_1(t)}d\widetilde{W}_1(t), \\ dY_2(t) &= -\lambda_2 Y_2(t)dt + \sigma_{21}\sqrt{Y_1(t)}d\widetilde{W}_1(t) \\ &\quad + \sqrt{\alpha + \beta Y_1(t)}d\widetilde{W}_2(t). \end{aligned}$$

We assume $\mu \geq 0, \lambda_1 > 0, \lambda_2 > 0, \alpha \geq 0, \beta \geq 0$, and $\sigma_{21} \in \mathbb{R}$. The Brownian motions $\widetilde{W}_1(t)$ and $W_2(t)$ are independent. We assume $Y_1(0) \geq 0$, and we have $Y_1(t) \geq 0$ for all $t \geq 0$ almost surely. On the other hand, even if $Y_2(t)$ is positive, $Y_2(t)$ can take negative values for $t > 0$. The interest rate is defined by

$$R(t) = \delta_0 + \delta_1 Y_1(t) + \delta_2 Y_2(t)$$

In all the two-factor models, the bond price $B(t,T)$ can be obtained by using the Feymann-Kac formula and assuming that the solution is of the form $\exp(-A(t) - C_1(t)Y_{1,t} - C_2(t)Y_{2,t})$. In addition, the yield-to-maturity is again an affine function of the factors.

4.3 HJM models

The second approach is that we model the forward curve under $\widetilde{\mathbb{P}}$ directly, which are called HJM modes.

To begin with, recall that the forward rate $f(t,T)$ is given by $f(t,T) = -\frac{\partial \ln B(t,T)}{\partial T}$ and $r_t = f(t,t)$ Then we have

$$B(t,T) = \exp\left(-\int_t^T f(t,v)dv\right)$$

Suppose $f(t,T)$ satisfies the equation

$$f(t,T) = f(0,T) + \int_0^t \alpha(u,T)du + \int_0^t \sigma(u,T)dW_u.$$

or equivalently $df(t,T) = \alpha(t,T)dt + \sigma(t,T)dt$
By Ito's formula, let $g(t) = \int_t^T f(t,v)dv$

$$\begin{aligned}
B(t,T) &= \exp\left(-\int_t^T f(t,v)dv\right) = \exp(g(t)) \\
\Rightarrow dB(t,T) &= -B(t,T)dg(t) + \frac{1}{2}B(t,T)dg(t)dg(t).
\end{aligned}$$

By Ito product rule,

$$\begin{aligned}
dg(t) &= -f(t,t)dt + \int_t^T df(t,v)dv \\
&= -f(t,t)dt + \int_t^T \alpha(t,v)dtdv + \int_t^T \sigma(t,v)dW_t dv \\
&= -f(t,t)dt + \int_t^T \alpha(t,v)dvdt + \int_t^T \sigma(t,v)dvdW_t \\
&:= -r_t dt + \alpha^*(t,T)dt + \sigma^*(t,T)dW_t
\end{aligned}$$

where we define $\alpha^*(t,T) = \int_t^T \alpha(t,v)du$, $\sigma^*(t,T) = \int_t^T \sigma(t,v)dv$

$$\begin{aligned}
\Rightarrow dB(t,T) =& - B(t,T)\left(-r_t dt + \alpha^*(t,T)dt + \sigma^*(t,T)dW_t\right) + \frac{1}{2}B(t,T)\sigma^*(t,T)^2 dt \\
=& \left(r_t - \alpha^*(t,T) + \frac{1}{2}\left(\sigma^*\right)^2(t,T)\right) B(t,T)dt + \sigma^*(t,T)B(t,T)dW_t
\end{aligned}$$

For this model to exclude arbitrage, $D_t B(t,T)$ should be a martingale under $\widetilde{\mathbb{P}}$. By Ito formula,

$$\begin{aligned}
d\left[D_t B(t,T)\right] &= D_t dB(t,T) + B(t,T)D_t \\
&= D_t\left(r_t - \alpha^*(t,T) + \frac{1}{2}\left(\sigma^*\right)^2(t,T)\right)B(t,T)dt \\
&-D_t\sigma^*(t,T)B(t,T)dW_t - r_t DtB(t,T)\,dt \\
&= D_t B(t,T)\left[\left(-\alpha^*(t,T) + \frac{1}{2}\left(\sigma^*\right)^2(t,T)\right)dt - \sigma^*(t,T)dW_t\right] \\
&= -\sigma^*(t,T)D_t B(t,T)\left[\frac{\alpha^*(t,T) - \frac{1}{2}\left(\sigma^*\right)^2(t,T)}{\sigma^*(t,T)}dt + dW_t\right]
\end{aligned}$$

For $\tilde{\mathbb{P}}$ to exist, we have that

$$\theta_t = \frac{\alpha^*(t,T) - \frac{1}{2}\left(d^*\right)^2(t,T)}{\sigma^*(t,T)}$$

for some process $\{\theta_t\}$. Suppose $Z_T = \exp\left(-\frac{1}{2}\int_0^T \theta_s^2 ds - \int_0^T \theta_s dW_s\right)$ and $Z_t = E\left[Z_T \mid \mathcal{F}_t\right]$, consicler the new probability measure $\mathbb{P}$ defined by $\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} = Z_T$, then $\theta_t dt + dW_t = d\widetilde{W}_t$ is a standard Brownian motion under $\widetilde{p}$. Moreover, we can obtain the dynamics of $B(t,T), f(t,T)$ under $\widetilde{\mathbb{P}}$:

$$\begin{aligned}
&d\left[D_t B(t,T)\right] = -\sigma^*(t,T)D_t B(t,T)d\widetilde{W}_t \\
&dB(t,T) = r_t B(t,T)dt - \sigma^*(t,T)B(t,T)d\widetilde{W}_t \\
&df(t,T) = \left(\alpha(t,T) - \theta_t\sigma(t,T)\right)dt + \sigma(t,T)d\widetilde{W}_t
\end{aligned}$$

As we have

$$\theta_t\sigma^*(t,T) = \alpha^*(t,T) - \frac{1}{2}\left(\sigma^*\right)^2(t,T)$$

taking derivative w.r.t. T gives

$$\begin{aligned}
&\theta_t\sigma(t,T) = \alpha(t,T) - \sigma(t,T)\sigma(t,T) \\
\Rightarrow &df(t,T) = \sigma^*(t,T)\sigma(t,T)dt + \sigma(t,T)d\widetilde{W}_t.
\end{aligned}$$

5 Jump models

5.1 Review of Poisson processes

Up to now, all models are driven by BMs which has continuous paths. But they fail to capture many important features:

(1) Stock prices may shock

(2) Stock pries have fat-tail distributions, but BMs have thin tails

(3) Implied volatility may have smile/smirk.

One solution to capture these features is to use jump models.

Poisson processes are the processes that has independent increments with exponential arrival times $\tau \sim \exp(\lambda)$,i.e. $\mathbb{P}(\tau > t) = e^{-\lambda t}, \mathbb{E}[\tau] = \frac{1}{\lambda}$, which are memoryless: $(P(\tau > t+s \mid \tau > s) = P(\tau > t))$.

We also define the arrival time of the n-th jump $S_n := \sum_{i=1}^n i$, which has density $g_n(s) = \frac{(\lambda s)^{n-1}}{(n-1)!}\lambda e^{-\lambda s}$. The counting process $N_t = \max\{n \geqslant 0 : S_n \leqslant t\}$ has Poisson distribution with intensity λ (or pmf $P(N_t = k) = \frac{(\lambda t)^k}{k!}e^{-\lambda t}$ with variance λt)

Remark. The compensated Poisson process $M_t = N_t - \lambda t$ is a martingale.

Compound Poisson processes may have multiple arrivals at the same time. Let Y_i be i.i.d. random variables independent of N_t and $\beta = E[Y_i]$. The compound Poisson process is defined as $Q_t = \sum_{i=1}^{N_t} Y_i$. Similarly, $Q_t - \beta\lambda t$ is a martingale. Q_t also has stationary and independent increments, and its mgf is given by $e^{\lambda t(\varphi_Y(M)-1)}$, where ϕ_Y is the mgf of the random variable Y.

Conversely, if we decompose Q_t into Poisson processes according to the size of arrivals, they are also independent with the corresponding intensity (proof: compute the mgf).

5.2 Stochastic calculus for jump processes

Jump measures $J(ds \times dy) = \{\#$ of jumps at at with size $y\}$ Then $Q_t = \int_0^t \int_{\mathbb{R}+} yJ(ds \times dy)$. We also have the compensated jump measure $\tilde{J}(s,t] \times A) = J(s,t] \times A) - \lambda(t-s)G_Y(A)$, where G_Y is the distribution measure of the jump size Y.

We model jump diffusions X_t as

$$X_t = x_0 + \int_0^t \Gamma_s dW_s + \int_0^t \theta_s ds + J_t$$

where $X_t^c = X_0 + \int_0^t \Gamma_s dW_s + \int_0^t \theta_s ds$ is the continuous part and J_t is a pure jump process (piecewise constant and right-continuous), which we assume that J_t is also left-continunous at $t = 0$. Then for general adapted proves ϕ_t, the stochastic integral $\int_0^t \phi_s dX_s$ is defined as $\int_0^t \phi_s dX_s = \int_0^t \phi_s \Gamma_s dw_s + \int_0^t \phi_s \theta_s ds + \sum_{0<s\leq t} \phi_s \Delta J_s$ where $\Delta J_s = J_s - J_s$ - denotes the jump size of J_t at time $t = s$.

If $\{\phi_t\}$ is left-continuous, then $\int_0^t \phi_s dX_s$ is a martingale. For a counterexample when this condition does not hold, consider $X_s = M_s = N_s - \lambda s, \phi_s = \Delta N_s$, then

$$\int_0^t \phi_s dX_s = \int_0^t \phi_s dN_s - \lambda \int_0^t \phi_s ds = N_t - 0 = N_t$$

which is not a martingale.

We also have the quadratic variation and covariation: if

$$\begin{aligned} X_{1,t} &= X_{1,t}^c + J_{1,t} \\ X_{2,t} &= X_{2,t}^c + J_{2,t} \end{aligned}$$

Then

$$\begin{aligned} [X_1, X_1]_t &= [X_1^c, X_1^c]_t + \sum_{0<s<t} (\Delta J_{1,s})^2 \\ [X_1, X_2]_t &= [X_1^c, X_2^c]_t + \sum_{0<s<t} (\Delta J_{1,s})(\Delta J_{2,s}) \end{aligned}$$

which indicates that the continuous part and the pure jump part are orthogonal (in the $\mathcal{L}^2_{\mathrm{loc}}(\mathcal{M})$).

In differential form,

$$\begin{aligned} dX_{1,t}dX_{1,t} &= dX_{1,t}^c dX_{1,t}^c + (\Delta J_{1,t})^2 = dX_{1,t}^c dX_{1,t}^c + (\Delta X_{1,t})^2 \\ dX_{1,t}dX_{2,t} &= dX_{1,t}^c dX_{2,t}^c + \Delta X_{1,t}\Delta X_{2,t} \end{aligned}$$

This can be further generalized to if

$$\widetilde{X}_{1,t} = \int_0^t \phi_{1,s} dX_{1,s} \quad \tilde{X}_{2,t} = \int_0^t \phi_{2,s} dX_{2,s}$$

then we have

$$\left[\tilde{X}_1, \tilde{X}_2\right]_t = \int_0^t \phi_{1,s}\phi_{2,s} d\left[X_1^c, X_2^c\right]_s + \sum_{0<s<t} \phi_{1,s}\phi_{2,s}\Delta X_{1,s}\Delta X_{2,s}$$

in differential form,

$$d\tilde{X}_{1,t} d\tilde{X}_{2,t} = \phi_{1,t}\phi_{2,t} dX_{1,t}^c dX_{2,t}^c + \phi_{1,t}\phi_{2,t}\Delta X_{1,t}\Delta X_{2,t}$$

Similarly, the important Ito's formula claim that if f is C' and piecewise C^2, we have

$$\begin{aligned} f\left(X_t\right) = f\left(X_0\right) + \int_0^t f'\left(X_s\right) dX_s + \frac{1}{2}\int_0^t f''\left(X_s\right) dX_s^c dX_s^c \\ + \sum_{0<s\leq t}\left(f\left(X_s\right) - f\left(X_s^-\right)\right) \end{aligned}$$

and in differential form

$$df\left(X_t\right) = f'\left(X_t\right) dX_t^c + \frac{1}{2} f''\left(X_t\right) dX_t^c dX_t^c + \Delta\left[f\left(X_t\right)\right]$$

Using Ito's formula we can show that for any Brownian motion W_t, Poisson proves M_t they are independent and $[M, W]_t = 0$

We also have the 2D Ito formula: let $X_{1,t}, X_{2,t}$ t be jump diffusion and $f\left(t, x_1, x_2\right)$ be c^2, then

$$\begin{aligned} df\left(t, X_{1,t}, X_{2,t}\right) = & f_t dt + f_{x_1} dX_{1,t}^c + f_{x_2} dX_{2,t}^c \\ & + \frac{1}{2} f_{x_1 x_1} dX_{1,t}^c dX_{1,t}^c + f_{x_1 x_2} dX_{1,t}^c dX_{2,t}^c \\ & + \frac{1}{2} f_{x_2 x_2} dX_{2,t}^c dX_{2,t}^c + \Delta f\left(t, X_{1,t}, X_{2,t}\right) \end{aligned}$$

As a corollary, we have the Ito product rule:

$$\begin{aligned} d\left[X_{1,t} X_{2,t}\right] = & X_{1,t} dX_{2,t}^c + X_{2,t} dX_{1,t}^c + dX_{1,t}^c dX_{2,t}^c + \Delta\left(X_{1,t} X_{2,t}\right) \\ = & X_{1,t-} dX_{2,t}^c + X_{2,t-} dX_{1,t}^c + dX_{1,t}^c dX_{2,t}^c + \\ & X_{1,t-}\Delta X_{2,t} + X_{2,t-}\Delta X_{1,t} + \Delta X_{1,t}\Delta X_{2,t} \\ = & X_{1,t-} dX_{2,t} + X_{2,t-} dX_{1,t} + dX_{1,t} dX_{2,t} \end{aligned}$$

which follows the Doléans-Dade exponential of the jump diffusion X_t, $Z_t = \exp\left(X_t^c - \frac{1}{2}\left[X^c, X^c\right]_t\right)\prod_{0<s<t}\left(1 + \Delta X_s\right)$ to be the solution of the SDE $dZ_t = Z_t - dX_t$. Moreover, if X_t is a martingale, then Z_t is also a martingale.

If N_t is a Poisson process with rate λ and $Z_t = e^{(\lambda - \hat{\lambda})t}\left(\frac{\hat{\lambda}}{\lambda}\right)^{N_t}$, then the new probability measure $\tilde{\mathbb{P}}$ defined by $\frac{d\tilde{P}}{d\mathbb{P}} = Z_T$ satisfies N_t is a Poisson process with rate $\tilde{\lambda}$. (Proof: compute mgf)

Moreover, let $Q_t = \sum_{k=1}^K Y_k N_{k,t}$ be the compound Poisson proves with $N_{k,t}$ being independent Poisson processes with intensity $p_k\lambda$ and $\sum_{k=1}^K p_k = 1$. Then the new probability measure $\widetilde{\mathbb{P}}$ defined by $\frac{d\mathbb{P}}{d\mathbb{P}} = Z_T$, where

$$Z_t = \prod_{k=1}^k e^{\left(\lambda_k - \tilde{\lambda}_k\right)}\left(\frac{\tilde{\lambda}_k}{\lambda_k}\right)^{N_{k,t}}$$

and $\lambda_k = \lambda p_k$ satisfies that $Q_t = \sum_{k=1}^k Y_k\tilde{N}_{k_1 t}$ is a compound Poisson process with $\tilde{N}_{k,t}$ being independent and having intesity $\tilde{\lambda}_k$ under $\widetilde{\mathbb{P}}$

Moreover, if $Q_t = \sum_{k=1}^{N_t} Y_k$ is a compound Poisson process with random variable Y following distribution with density g. Suppose $\tilde{g}$ is anther density s.t. $g \neq 0$ whenever $\tilde{g} \neq 0$. Then if N_t is of intensity λ, then the new probability measure $\widetilde{\mathbb{P}}$ defined by $\frac{d\tilde{P}}{d\mathbb{P}} = Z_T$, where

$$Z_t = \exp\left(\lambda - \tilde{\lambda}\right)t\prod_{i=1}^{N(t)}\frac{\tilde{\lambda}\tilde{g}(Y_i)}{\lambda g(Y_i)}$$

satisfies that Q_t is a compound poisson process with intensity $\hat{\lambda}$ and size of jumps follows distribution $\hat{f}$.

Combining the result in the continuous case, if

$$\begin{aligned} Z_{1,t}\exp\left(\lambda - \tilde{\lambda}\right)t\prod_{i=1}^{N(t)}\frac{\tilde{\lambda}\tilde{g}(Y_i)}{\lambda g(Y_i)} \\ Z_{2,t} = \exp\left(-\frac{1}{2}\int_0^t\theta_s^2 ds - \int_0^t\theta_s dw_s\right) \end{aligned}$$

and $\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}} = Z_{1,T} Z_{2,T}$, then not only $Q_t \sim$ compound Poisson $(\lambda, g(\cdot))$ under $\mathbb{P}$ would be transformed into Q_t comp Poi $(\tilde{\lambda}, \tilde{g}(1))$ under $\widetilde{\mathbb{P}}$, but also $d\widetilde{W}_t = dW_t + \theta_t dt$ would be a BM under $\widetilde{\mathbb{P}}$.

Now we apply these theories to price assets. Suppose Q_t is a impound Poisson process with intensity λ and size density $g(\cdot)$, and the expected number of size of jump is β. Let W_t be a BM. Then a risky asset S_t follows the dynamics

$$\begin{aligned} dS_t &= \alpha S_t dt + \sigma S_t dW_t + S_{t-} d\left[Q_t - \beta\lambda t\right] \\ &= (\alpha - \beta\lambda) S_t dt + \sigma S_t dW_t + S_{t-} dQ_t - \end{aligned}$$

By Ito formula,

$$S_t = S_0 \exp\left(\left(\alpha - \beta\lambda - \frac{1}{2}\sigma^2\right) t + \sigma W_t\right) \prod_{i=1}^{N(t)} (Y_i + 1)$$

We hope that this market model excludes arbitrage, which means that we hope to find a risk-neutral probability $\widetilde{\mathbb{P}}$ s.t.

$$dS_t = r S_t dt + \sigma S_t dw_t + S_{t-} d\left[Q_t - \widetilde{\beta}\widetilde{\lambda} t\right]$$

where Q_t is a compound Poisson process with intensity $\tilde{\lambda}$ and expected size $\tilde{\beta}$. Rearranging the terms we have

$$dS_t = (r - \tilde{\beta}\tilde{\lambda}) S_t dt + \sigma S_t \left(dW_t + \theta_t dt\right) + S_{t-} dQ_t$$

for some process (θ_t) in the change-of-measure formula above.
Hence we only require

$$r - \tilde{\beta}\tilde{\lambda} + \sigma\theta_t = \alpha - \beta\lambda$$

Actually we have several ways t choose the parameters:
(1) We keep the jump process and have $\beta = \tilde{\beta}, \lambda = \tilde{\lambda}$, then $\theta = \frac{\alpha - r}{\sigma}$
(2) We keep the BM and have $\theta = 0$, keep the jump distribution and have $\beta = \tilde{\beta}$, then $\gamma - \beta\tilde{\lambda} = \alpha - \beta\lambda \Rightarrow \tilde{\lambda} = \lambda + \frac{\gamma - \alpha}{\beta}$
(3) We change the distribution of jump sizes from $g(\cdot)$ to $\hat{g}(\cdot)$.

5.3 Pricing derivatives driven by jump processes

Now we price the derivatives. The fair price of the European call is

$$\begin{aligned} & \widetilde{\mathbb{E}}\left[e^{-r(T-t)} \max\left(S_T - K, 0\right) \mid \mathcal{F}_t\right] \\ =& \widetilde{\mathbb{E}}\left[e^{-r(T-t)} \max\left(S_t e^{\sigma\left(\tilde{W}_T - \tilde{W}_t\right) + \left(r - \hat{\beta}\hat{\lambda} - \frac{1}{2}\sigma^2\right)(T-t)} \prod_{i=N_{t+1}}^{N_T} (1 + Y_i) - K, 0\right) \mid \mathcal{F}_t\right] \\ =& \widetilde{\mathbb{E}}\left[\widetilde{\mathbb{E}}\left[e^{-r(T-t)} \max\left(R S_t e^{\sigma\left(\tilde{W}_T - \tilde{w}_t\right) + \left(r - \frac{1}{2}\sigma^2\right)(T-t)} - K, 0\right) \mid R\right] \mid \mathcal{F}_t\right] \end{aligned}$$

where $R = e^{-\hat{\beta}\hat{\lambda}(T-t)} \sum_{i=N_{t+1}}^{N_T} (1 + Y_i)$. This is just

$$\begin{aligned} & \widetilde{\mathbb{E}}\left[BS_{\text{call }}\left(T - t, RS_t\right) \mid \mathcal{F}_t\right] \\ =& \widetilde{\mathbb{E}}\left[\widetilde{\mathbb{E}}\left[BS_{\text{call }}\left(T - t, RS_t\right) \mid N_T - N_t, \mathcal{F}_t\right] \mid \mathcal{F}_t\right] \\ =& \tilde{\mathbb{E}}\left[\sum_{j=0}^{\infty} e^{-\tilde{\lambda}(T-t)} \frac{[\tilde{\lambda}(T-t)]^j}{j!} BS_{\text{call}}\left(T - t, e^{-\tilde{\beta}\tilde{\lambda}(T-t)} \prod_{i=1}^{j} (1 + Y_i)\right) \mid \mathcal{F}_t\right] \\ =& \sum_{j=0}^{\infty} e^{-\tilde{\lambda}(T-t)} \frac{[\tilde{\lambda}(T-t)]^j}{j!} \tilde{\mathbb{E}}\left[BS_{\text{call}}\left(T - t, e^{-\tilde{\beta}\tilde{\lambda}(T-t)} \prod_{i=1}^{j} (1 + Y_i)\right) \mid \mathcal{F}_t\right] \end{aligned}$$

where $BS_{\text{call }}\left(T, S_0\right)$ is the price of an European call option under the BS model with maturity T and initial stock price S_0. Now we provide a partial-integral-differential approach for the option price. By argument above, $\exists$ a function $c(t, x)$ s.t. it is the fair for the European call at time t with underlying asset prise x. Hence by Ito formula,

$$d\left[e^{-rt} c\left(t, S_t\right)\right] = -r e^{-rt} c\left(t, S_t\right) dt + e^{-rt} dc\left(t, S_t\right)$$

and we have

$$\begin{aligned} dc\left(t, S_t\right) =& c_t\left(t, S_t\right) dt + c_x\left(t, S_t\right) dS_t^c + \frac{1}{2} c_{xx}\left(t, S_t\right) dS_t^c dS_t^c \\ & + \Delta c\left(t, S_t\right) \\ =& c_t\left(t, S_t\right) dt + c_x\left(t, S_t\right)(r - \tilde{\beta}\tilde{\lambda}) dt + \frac{1}{2}\sigma^2 S_t^2 c_{xx}\left(t, S_t\right) dt \\ & + \sigma c_x\left(t, S_t\right) S_t d\tilde{W}_t + \Delta c\left(t, S_t\right) \end{aligned}$$

where the jumps are given by

$$\begin{aligned}\Delta c\left(t, S_t\right) =& c\left(t, S_t\right)-c\left(t, S_t\right) \\ =& \int_{-1}^{\infty}\left[c\left(t, S_{t-}(1+y)\right)-c\left(t, S_{t-}\right)\right] J(d y \times d t) \\ =& \int_{-1}^{\infty}\left[c\left(t, S_{t-}(1+y)\right)-c\left(t, S_{t-}\right)\right] \tilde{J}(d y \times d t)+ \\ & \int_{-1}^{\infty}\left[c\left(t, S_{t-}(1+y)\right)-c\left(t, S_{t-}\right)\right] \tilde{\lambda} \tilde{g}(y) d y d t\end{aligned}$$

where $J(dy \times dt)$ is the jump measure for the compound Poisson process and $\tilde{J}(dy \times dt) = J(dy \times dt) - \tilde{\lambda}\tilde{g}(y)dydt$ is the compensated measure s.t. the integration is a martingale.
Hence the dynamics of $e^{-rt}c\left(t, S_t\right)$ is given by

$$\begin{aligned}d\left[e^{-rt} c\left(t, S_t\right)\right] =& e^{-2t}\left[\left(c_t-\tilde{\beta} \tilde{\lambda} c_x+\frac{1}{2} d^2 S_t^2 c_{xx}\right) d t+\sigma c_x d \widetilde{W}_t\right) \\ & +\int_{-1}^{\infty}\left[c\left(t, S_{t-}(1+y)\right)-c\left(t, S_{t-}\right)\right] \tilde{J}(d y \times d t)+ \\ & \int_{-1}^{\infty}\left[c\left(t, S_{t-}(1+y)\right)-c\left(t, S_t\right)\right] \tilde{\lambda} g(y) d y d t\end{aligned}$$

Setting the dt term to zero we have

$$c_t-\tilde{\beta} \tilde{\lambda} c_x+\frac{1}{2} \sigma^2 x^2 c_{xx}+\tilde{\lambda} \int_{-1}^{\infty}[c(t, x(1+y))-c(t, x)] g(y) d y=0$$

with terminal condition $c(T, x) = (x-k)^+$. Now suppose that we would like to hedge this option by an SFTS $\{X_t\}$ that holds Γ_t shares of risky assets, then

$$\begin{aligned}d\left[e^{-rt} X_t\right] &= -r e^{-et} X_t d t+e^{-rt} d X_t \\ &= e^{-rt} \Gamma_t \sigma S_t d \widetilde{W}_t+e^{-rt} \Gamma_{t-} S_{t-} \int_1^{\infty} y \tilde{J}(d y \times d t)\end{aligned}$$

and also

$$d\left[e^{-rt} c\left(t, S_t\right)\right] = e^{-rt} \sigma_t S_t c_x d \widetilde{W}_t+e^{-rt} \int_{-1}^{\infty}\left[c\left(t, S_{t-}(1+y)\right)-c\left(t, S_{t-}\right)\right] \hat{J}(d y \times d t)$$

Hence if we set $c_x = \Gamma_t$ to eliminate the $d\widetilde{W}_t$ term,

$$\begin{aligned}d\left[e^{-rt}\left(X_t-c\left(t, S_t\right)\right]\right. = e^{-rt} \int_{-1}^{\infty} & \left[c\left(t, S_t-\right)+y S_{t-} c_x\left(t, S_t\right)\right. \\ & \left.-c\left(t, S_{t-}(1+y)\right] d \tilde{J}(d y \times d t)\right.\end{aligned}$$

If there is no jumps, then

$$\int_{-1}^{\infty} J(d y \times d t)=0 \Rightarrow \int_{-1}^{\infty} \tilde{J}(d y \times d t)=\int_{-1}^{\infty} J(d y \times d t)-\int_{-1}^{\infty} \tilde{\lambda} g(y) d y d t \leqslant 0$$

hence there is a trend to go down in $e^{-rt}(xt - c(t, St))$, the hedging portfolio underperforms.
If there is a jump of size y, then as $c(t, x)$ is convex,

$$c\left(t, S_{t-}(1+y)\right)-c\left(t, S_{t-}\right)<y S_{t-} c_x\left(t, S_{t-}\right)$$

$\Rightarrow d\left[e^{-rt}(X_t - c(t, St)\right]$ has the trend to go up.
In this case, the hedging portfolio outperforms.