THE STRONG MARKOV PROPERTY APPLIED TO OPTION PRICING PART I

MASAHIKO EGAMI

Graduate School of Economics, Kyoto University, Sakyo-ku, Kyoto, 606-8501, Japan.

ABSTRACT. In pricing and hedging derivatives, one uses stochastic processes to represent movement of underlying assets and other important state variables (such as volatility). In this report, we collect several techniques from stochastic analysis which we believe are useful for financial modeling, derivative pricing, and parameter estimation. A key element is the strong Markov property that helps to streamline seemingly complex equations and to tremendously reduce variance in numerical simulations.

1. INTRODUCTION

In pricing and hedging derivatives, one uses stochastic processes to represent movement of underlying assets (stock price, index, interest rate) and other state variables such as volatility. Perhaps, the most popular is geometric Brownian motion, which is explained in almost all financial engineering books. A brief review is also in *Futures and Options Report* (Egami (2022)).

Let us consider a set of all the possibilities of future outcomes Ω and denote our stochastic process by $X = \{X_t; t \geq 0\}$ with the initial value x. In this report, we assume that the process X is *right continuous*. Given a future scenario $\omega \in \Omega$ and a fixed time t, the stochastic process X takes a value $X_t(\omega)$ in some space (such as $\mathbb{R}^d$). Given this ω, X keeps taking values as time goes and a history of its behavior is recorded. This is called the *filtration* generated by X and denoted by $\mathscr{F} = \{\mathscr{F}_t; t \geq 0\}$. Intuitively, $\mathscr{F}_t$ is the accumulated information of X up to time t. While it is possible to enlarge the filtration including more information than the behavior of X, we confine to the above-defined $\mathscr{F}$ in this report.

If $\{X_t, \mathscr{F}_t, t \geq 0\}$ is a *Markov* process, it has the property for a certain set A in $\mathbb{R}^d$

$$\mathbb{P}_x(X_{s+t} \in A | \mathscr{F}_s) = \mathbb{P}_x(X_{s+t} \in A | X_s), \tag{1.1}$$

which roughly means the following. Suppose that the process X started at the value of x at time zero and we are now at time s and wish to estimate the value of X at $s+t$, more specifically, the probability that the value is in the set A. The Markov property says that, for this purpose, we do not have to bring all the records of X up to time s, because looking at the sole value of X at time s is equally effective. This property saves us a vast amount of time and effort (e.g. in simulation) since we do not have to consider through what path X has come to the value X_s. For another example, take a random variable Y: at time s we do not know what value is, but we may estimate its value by looking into future behavior of X. One could imagine that X represents a stock price and Y is the value of an option written on this stock. Then if X is a Markov process, we have

$$\mathbb{E}_x[Y|\mathscr{F}_s] = \mathbb{E}_x[Y|X_s]. \tag{1.2}$$

A *strong Markov* process does better than this. The above (1.1) holds even if the usual time t is replaced by a random time U of a special kind (called *stopping time*). Since it is random, it depends on a future scenario ω and we do not know a priori when a random time $U(\omega)$ occurs. A random time U is called a stopping time of $\mathscr{F}$ provided that we can tell whether or not U has happened up to time t by looking at the record of X up to time t. That is, $\{\omega : U(\omega) \leq t\} \in \mathscr{F}_t$ for each t. If $\{X_t, \mathscr{F}_t, t \geq 0\}$ is a *strong Markov* process, it has the property for a certain set A in $\mathbb{R}^d$

$$\mathbb{P}_x(X_{U+t} \in A|\mathscr{F}_U) = \mathbb{P}_x(X_{U+t} \in A|X_U), \tag{1.3}$$

where $\mathscr{F}_U$ is the record of X up to the stopping time U.

Frequently used stochastic processes in financial applications such as diffusions and Lévy processes are strongly Markov.

2. Pricing a Barrier Option

A motivating example is the valuation of barrier options. Suppose that a stock is denoted by $S = \{S_t; t \geq 0\}$. A *down-and-in* call has the payoff function f at maturity T with strike K

$$f(x,\omega) := (x-K)^+ \mathbb{1}_D(\omega) = \max(0, x-K)\mathbb{1}_D(\omega)$$

where, denoting the predetermined barrier by b, the indicator $\mathbb{1}_D$ means

$$\mathbb{1}_D(\omega) := \begin{cases} 1, & \text{if } S_u(\omega) \leq b \text{ for some } 0 \leq u \leq T, \\ 0, & \text{otherwise.} \end{cases}$$

In other words, the option has a positive value only if the stock price falls below the level b before maturity T. Let us define the random time H_b by

$$H_b := \min\{t \geq 0 : S_t \leq b\}, \tag{2.1}$$

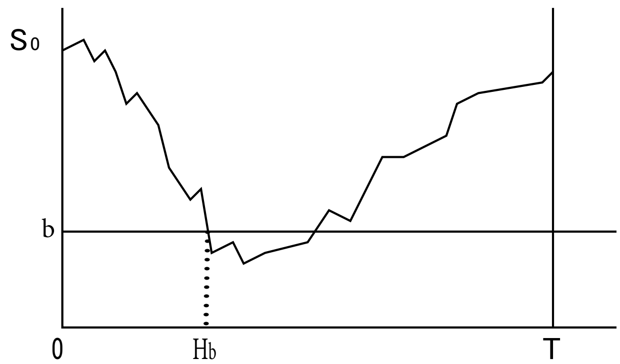
which is the first time that S hits the level b. See Figure 1 for a path where the option has positive value. This H_b is obviously a stopping time since one can tell whether this event has occurred up to time t only by looking at the stock price up to time t. Let the initial stock price be $S_0 = x(> b)$ and the risk-free rate r. The time zero price $P(0,x)$ of the down-and-in call is computed by the well-known risk-neutral pricing formula

$$P(0,x) = \mathbb{E}_x^*[e^{-rT}(S_T(\omega) - K)^+ \mathbb{1}_D(\omega)]$$

where the expectation $\mathbb{E}_x^*[\cdot]$ indicates that the stock price starts with $S_0 = x$ and we are taking expectations under the risk-neutral probability measure.

It is often a good strategy to condition on the information up to the stopping time as we shall show now. Let us define the price of *plain vanilla* call at time t written on the same stock S by $Q(t, S_t)$ with

Fig. 1. DOWN-AND-IN OPTION: A sample path where a down-and-in option has positive value



the starting date t and stock price at that time S_t, so that the remaining time to maturity is $T-t$. That is, $Q(t,S_t)=\mathbb{E}_x^*[e^{-r(T-t)}(S_T-K)^+|S_t]$. With this notation, $P(0,x)$ above becomes

$$\begin{aligned}
P(0,x) &= \mathbb{E}_x^*[\mathbb{E}_x^*[e^{-rH_b}\mathbb{1}_D(S_T-K)^+|\mathscr{F}_{H_b}]] \\
&= \mathbb{E}_x^*[e^{-rH_b}\mathbb{1}_D\mathbb{E}_x^*[e^{-r(T-H_b)}(S_T-K)^+|\mathscr{F}_{H_b}]] \\
&= \mathbb{E}_x^*[e^{-rH_b}\mathbb{1}_D\mathbb{E}_x^*[e^{-r(T-H_b)}(S_T-K)^+|S_{H_b}]] \\
&= \mathbb{E}_x^*[e^{-rH_b}\mathbb{1}_D Q(H_b,S_{H_b})], \qquad (2.2)
\end{aligned}$$

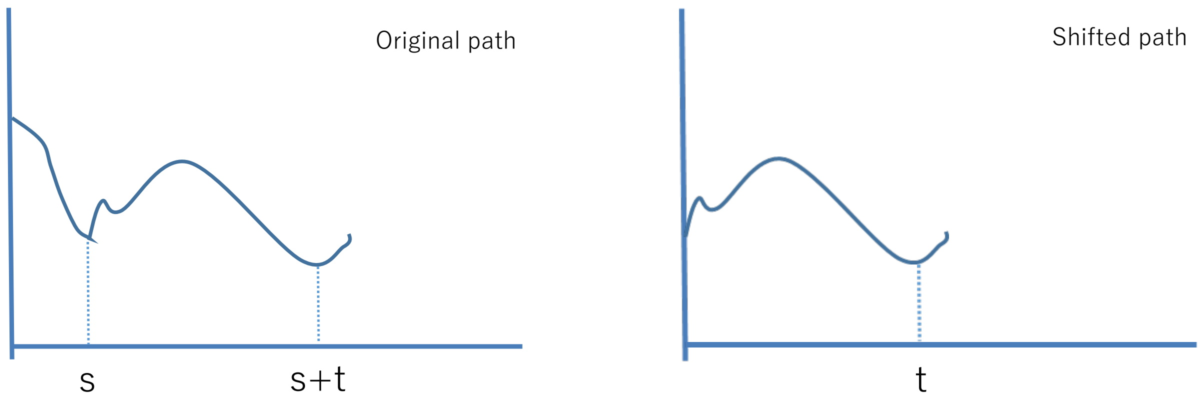
where the first line is the tower property, the second is *taking out what is known* at time H_b (see [Section II.41, Rogers and Williams (1994)]) and the third is due to the strong Markov property. At time H_b, it becomes to our knowledge that the event D has just occurred. This is why $\mathbb{1}_D$ is taken out from the inner expectation in the second line. All we need to know from the record of S is its price at time H_b. For the last line, we just apply the above definition of $Q(\cdot,\cdot)$ by reading t and S_t as H_b and S_{H_b}, respectively.

If S follows a geometric Brownian motion, then $S_{H_b}=b$, the density of the first passage time $\mathbb{P}_x(H_b\in \mathrm{d}t)$ is available. See page 628 in Borodin and Salminen (2002). Alternatively, one can use the Laplace transform of the distribution of H_b (see the next section as well as the quoted book) to generate random samples by the method proposed in Ridout (2009). In either case, we only need to evaluate or simulate H_b. Finally, $Q(H_b,b)$ is computed by the Black-Scholes formula with the initial stock price b and time to maturity $T-H_b$ since it is merely the plain vanilla call price.

3. THE SHIFT OPERATOR, α-POTENTIAL, AND DYNKIN'S FORMULA

We have seen that the strong Markov property can make seemingly complicated problems easier. For concrete computations and for practical applications, let $X=(X_t,t\geq 0)$ be a diffusion process taking values on $\mathbb{R}^d$. The interested readers may see Karlin and Taylor (1981) for detailed information of diffusions, but it suffices to consider geometric Brownian motion for this report. The shift operator

Fig. 2. THE SHIFT OPERATOR: The left panel is $X_t(\omega)$ and the right panel is $X_t(\theta_s(\omega))$. Note that the path on the right panel starts at the point corresponding to (s, X_s) on the left panel. The path on the left from time zero to $t-$ (just prior to time t) is irrelevant information in drawing the path on the right panel.



$\theta_s : \Omega \to \Omega, s \geq 0$ further facilitates rigorous argument:

$$X_{s+t}(\omega) = X_t(\theta_s(\omega)) = (X_t \circ \theta_s)(\omega) \quad \text{for each } s, t \geq 0. \tag{3.1}$$

Let us examine the diagram in Figure 2. The shift operator *shifts* the time (in a future scenario) by the length of s and makes the point s as a new "origin". From the new origin s, the time elapses by the length of t. If we use this shift operator, (1.1) can be written

$$\mathbb{P}_x(X_{s+t}(\omega) \in A | \mathscr{F}_s) = \mathbb{P}_x\{(X_t \circ \theta_s)(\omega) \in A | \mathscr{F}_s\} = \mathbb{P}_{X_s}[X_t(\omega) \in A],$$

that is, the process X starts anew from time s at the position X_s. Similarly, (1.2) is written for the derivative Y written on the underlying asset X

$$\mathbb{E}_x[Y(\omega)|\mathscr{F}_s] = \mathbb{E}_x[(Y \circ \theta_s)(\omega)|\mathscr{F}_s] = \mathbb{E}_{X_s}[Y(\omega)] \tag{3.2}$$

For simpler exposition, we shall omit the argument ω and write X_t instead of $X_t(\omega)$ from this point on. Now let f be a bounded continuous function and $\alpha > 0$, the discount rate. We introduce the function $U^\alpha f$, which is called the α-potential of f and is defined by

$$U^\alpha f(x) := \mathbb{E}_x\left[\int_0^\infty e^{-\alpha t} f(X_t) \mathrm{d}t\right].$$

This is very important and useful for financial modelling. For example, if f and X denote a dividend function and stock price at time t, respectively, then $U^\alpha f$ is the present value (discounted at rate α) of the total amount of dividend accumulated until time infinity.

Let T be a stopping time and X a strong Markov process. Let us imagine that we observe the stock price X up to the stopping time T and split the $U^{\alpha}f(x)$ before and after the stopping time T,

$$U^{\alpha}f(x) = \mathbb{E}_x\left[\int_0^T e^{-\alpha t}f(X_t)\mathrm{d}t\right] + \mathbb{E}_x\left[\int_T^{\infty} e^{-\alpha t}f(X_t)\mathrm{d}t\right].$$

Recall that x is the initial level of X at time zero. We then have

$$\mathbb{E}_x\left[\int_T^{\infty} e^{-\alpha t}f(X_t)\mathrm{d}t\right] = \mathbb{E}_x\left[\mathbb{E}_x\left\{\int_T^{\infty} e^{-\alpha t}f(X_t)\mathrm{d}t\middle|\mathscr{F}_T\right\}\right] = \mathbb{E}_x[e^{-\alpha T}U^{\alpha}f(X_T)],$$

where we use the tower property for the first equality and the strong Markov property, with a help of the shift operator, for the second. We refer the interested reader to Appendix A (to be posted in Part II) for a derivation. Hence we have obtained

$$U^{\alpha}f(x) = \mathbb{E}_x\left[\int_0^T e^{-\alpha t}f(X_t)\mathrm{d}t\right] + \mathbb{E}_x[e^{-\alpha T}U^{\alpha}f(X_T)]. \tag{3.3}$$

The meaning of the term $\mathbb{E}_x[e^{-\alpha T}U^{\alpha}f(X_T)]$ in (3.3) is intuitively clear: Given the information up to time T, $U^{\alpha}f(X_T)$ is the amount of dividend received from time T to infinity, but by starting with the stock price at X_T (instead of x). Since this amount $U^{\alpha}f(X_T)$ is calculated at time T, we need to discount it from time T to time zero for making it the present value at time zero. This is done by multiplying $e^{-\alpha T}$.

Since one further step shall bring (3.3) to the celebrated *Dynkin's formula*, let us introduce another object of diffusions. The infinitesimal generator $\mathfrak{G}$ of a diffusion X is defined by

$$\mathfrak{G}g(x) := \lim_{t\downarrow 0}\frac{\mathbb{E}_x[g(X_t)] - g(x)}{t} \tag{3.4}$$

for a suitable function g. This is how much the value $g(X_t)$ would change if an infinitely small amount of time has elapsed from time zero. This generator is closely related to the above-mentioned α-potential in an interesting way

$$U^{\alpha}(\alpha - \mathfrak{G})g = g, \tag{3.5}$$

that is, $(\alpha - \mathfrak{G})$ and U^{α} are inverse to each other. For a proof, see [Section III.4, Rogers and Williams (1994)]. Now let us set $f = (\alpha - \mathfrak{G})g$ in (3.3) to obtain

$$U^{\alpha}(\alpha - \mathfrak{G})g(x) = \mathbb{E}_x\left[\int_0^T e^{-\alpha t}(\alpha - \mathfrak{G})g(X_t)\mathrm{d}t\right] + \mathbb{E}_x[e^{-\alpha T}U^{\alpha}(\alpha - \mathfrak{G})g(X_T)].$$

But with (3.5), it reduces to

$$g(x) = \mathbb{E}_x\left[\int_0^T e^{-\alpha t}(\alpha - \mathfrak{G})g(X_t)\mathrm{d}t\right] + \mathbb{E}_x[e^{-\alpha T}g(X_T)],$$

which, after rearrangement, is called Dynkin's formula in regard to a stopping time T

$$\mathbb{E}_x[e^{-\alpha T}g(X_T)] - g(x) = \mathbb{E}_x\left[\int_0^T e^{-\alpha t}(\mathfrak{G} - \alpha)g(X_t)\mathrm{d}t\right]. \tag{3.6}$$

Let us apply (3.6) for an explicit computation of $\mathbb{E}_x[e^{-\alpha H_b}]$ where $H_b = \min\{t \geq 0 : S_t \leq b\}$ as in (2.2). Let X be a geometric Brownian motion with drift parameter $\mu \in \mathbb{R}$ and diffusion parameter $\sigma > 0$

$$X_t = X_0 \cdot e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t} = x \cdot e^{\sigma^2 \nu t + \sigma W_t},$$

where $\nu = \frac{\mu}{\sigma^2} - \frac{1}{2}$. The infinitesimal generator for a linear diffusion $\mathrm{d}X_t = \mu(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t$ is

$$\mathfrak{G}g(x) = \frac{1}{2}\sigma^2(x)\frac{\mathrm{d}^2}{\mathrm{d}x^2}g(x) + \mu(x)\frac{\mathrm{d}}{\mathrm{d}x}g(x), \tag{3.7}$$

which is essentially the dt term in the Itô formula. In the geometric Brownian motion case, $\mu(x) = \mu x$ and $\sigma(x) = \sigma x$, so that $\mathfrak{G}g(x) = \frac{1}{2}\sigma^2 x^2 \frac{\mathrm{d}^2}{\mathrm{d}x^2}g(x) + \mu x \frac{\mathrm{d}}{\mathrm{d}x}g(x)$. The two solutions to the equation $(\mathfrak{G} - \alpha)g(x) = 0$ are

$$g(x) = x^{\sqrt{\nu^2 + \frac{2\alpha}{\sigma^2}} - \nu} \quad \text{and} \quad g(x) = x^{-\sqrt{\nu^2 + \frac{2\alpha}{\sigma^2}} - \nu}.$$

(Try a function of the form $x^\beta, \beta \in \mathbb{R}$ and compute $(\mathfrak{G} - \alpha)x^\beta = 0$. The exponent β is determined by the requirement that the latter equation be held for *any* $x \in \mathbb{R}$.)

Recall that the starting stock price x is greater than the threshold level b and set $T = H_b$ with

$$g(x) := x^{-\sqrt{\nu^2 + \frac{2\alpha}{\sigma^2}} - \nu} \quad \text{for } b \leq x < \infty$$

in (3.6). While there are two solutions of $(\mathfrak{G} - \alpha)g(x) = 0$, we need to choose this solution since the other one becomes unbounded as $x \to \infty$. Note that $X_{H_b} = b$ and that from time zero to H_b, X_t is in the interval of (b, ∞). On (b, ∞), $(\mathfrak{G} - \alpha)g = 0$ by the construction of $g(\cdot)$. Therefore, (3.6) becomes

$$\mathbb{E}_x[e^{-\alpha H_b} g(b)] - g(x) = 0$$

on the set $\{\omega : H_b(\omega) < \infty\}$. Hence we have established

$$\mathbb{E}_x[e^{-\alpha H_b}] = \frac{g(x)}{g(b)} = \left(\frac{x}{b}\right)^{-\sqrt{\nu^2 + \frac{2\alpha}{\sigma^2}} - \nu}$$

on the set $\{\omega : H_b(\omega) < \infty\}$. This is a powerful method to compute various functionals explicitly provided that we know the generator of specific interest.

References

Borodin, A. N., Salminen, P., 2002. Handbook of Brownian Motion - Facts and Formulae, 2nd Edition. Birkhäuser, 2015 corrected reprint.

Egami, M., 2022. A financial engineering view of drawdowns in the stock market part I. Futures and Options Report No. 1, Osaka Exchange.

Karlin, S., Taylor, H. M., 1981. A Second Course in Stochastic Processes. Academic Press.

Ridout, M., 2009. Generating random numbers from a distribution specified by its Laplace transform. Stat Comput 19, 439–450.

Rogers, L., Williams, D., 1994. Diffusions, Markov processes and martingales, Vol.1: 2nd Edition. Cambridge University Press, Cambridge.