Chapter 6

Approximating the underlying stochastic process – The tree methodology

6.1 Introduction

In chapter 2 section 2.6 we discussed the Black-Scholes-Merton model where the stock process follows a geometric Brownian motion:

$$dS_t = \mu S_t dt + \sigma S_t dB_t. \tag{6.1}$$

In this chapter we will learn tree approximation methods for the stock process in (6.1). We refer to [38] for pseudo-code in the basic trees described here.

The stochastic process in (6.1) is a particular case of a Markov process. A tree approximation is nothing more than a Markov chain approximation for a Markov process. Therefore, for generalization purposes it is important to understand how we create trees that approximate the diffusion process in equation (6.1).

6.2 Markov Process

Markov processes were briefly introduced in chapter 1 section 1.4 of this book. Here we reiterate the definitions.

Let $\{X_t\}_{t\geq 0}$ be a process on the space $(\Omega, \mathcal{F}, \mathcal{P})$ with values in E and let $\mathcal{G}_t$ a filtration such that X_t is adapted with respect to this filtration. In the most basic case the filtration is generated by the stochastic process itself: $\mathcal{G}_t = \mathcal{F}_t^x =$

$\sigma(X_s : s \leq t)$. The process X_t is a Markov process if and only if:

$$\mathcal{P}(X_{t+s} \in \Gamma | \mathcal{G}_t) = \mathcal{P}(X_{t+s} \in \Gamma | X_t), \forall s, t \geq 0 \text{ and } \Gamma \in \mathcal{B}(E). \tag{6.2}$$

the collection $\mathcal{B}(E)$ denote the Borel sets of E. Essentially, the definition says in order to decide where the process goes next knowing the current state is the same as knowing the entire set of past states. The process in some sense does not have memory of past states.

Equation (6.2) is equivalent to

$$E\left[f(X_{t+s})|\mathcal{F}_t^X\right] = E\left[f(X_{t+s})|X_t\right],$$

$\forall$ f Borel measurable functions on E.

6.2.1 Transition function

Since knowing the current state is enough for determining the future, the random variable $X_t|X_s$ determines the process trajectories. This random variable has a distribution and if the density of the distribution exists this is called a transition function. Specifically, the random variable $X_t|X_s = x$, has a density which is typically denoted $p(s,t,x,y)$:

$$\mathcal{P}(s,t,x,\Gamma) = \int_\Gamma p(s,t,x,y)dy.$$

In the special case where the Markov process is homogeneous (that is $X_{t+s}|X_s$ has the same distribution as $X_t|X_0$ for all s and t) the transition function simplifies to $p(s,t,x,y) = p(t-s,x,y)$. This special case is important since all diffusion equations of the type:

$$dS_t = \mu(S_t)dt + \sigma(S_t)dB_t. \tag{6.3}$$

where μ and σ are functions that do not have a t argument, will produce a homogeneous Markov process. All the models used in practice in Finance are of this type.

Formally, a function $\mathcal{P}(t,x,\Gamma)$ defined on $[0,\infty) \times E \times \mathcal{B}(E)$ is a (time homogeneous) transition function if:

1. $\mathcal{P}(t,x,\cdot)$ is a probability measure on $(E,\mathcal{B}(E))$.
2. $\mathcal{P}(0,x,\cdot) = \delta_x$ (Dirac measure).
3. $\mathcal{P}(\cdot,\cdot,\Gamma)$ is a Borel measurable function on $[0,\infty) \times E$.
4. The function satisfies the Chapman-Kolmogorov equation $\mathcal{P}(t+s,x,\Gamma) = \int \mathcal{P}(s,y,\Gamma)\mathcal{P}(t,x,dy)$, $s,t \geq 0, x \in E$ and $\Gamma \in \mathcal{B}(E)$.

The connection is the following. A transition function is the transition function for a time homogeneous Markov process X_t if and only if:

$$\mathcal{P}(X_{t+s} \in \Gamma | \mathcal{F}_t^x) = \mathcal{P}(s, X_t, \Gamma),$$

that means it actually expresses the probability that the process goes into a set Γ at time $t+s$ given that it was at X_t at time t. Furthermore, all these properties may be expressed in terms of the density function $p(t,x,y)$

$$\mathcal{P}(t,x,\Gamma) = \int_\Gamma p(t,x,y)dy,$$

if the density function exists.

The next theorem stated without proof insures that our discussion of transition functions is highly relevant.

Theorem 6.2.1. *Any Markov process has a transition function. If the metric space (E,r) is complete and separable then for any transition function there exists a unique Markov process with that transition function.*

If the Markov process is defined on $\mathbb{R}$ then the theorem is valid and there is a one to one equivalence between Markov processes and their transition functions. However, $\mathbb{R}^2$ and higher orders are not separable spaces. So, there may be multiple Markov processes with the same transition function. This is actually useful if one wants to generate a transition distribution, one may use a simpler process to generate the distribution of complicate ones.

6.2.2 Strong Markov process

A Markov process $\{X_t\}_{t\geq 0}$ with respect to $\{\mathcal{G}_t\}_{t\geq 0}$ is called a strong Markov process at ζ if:

$$\mathcal{P}(X_{t+\zeta} \in \Gamma | \mathcal{G}_\zeta) = \mathcal{P}(t, X_\zeta, \Gamma) \tag{6.4}$$

where ζ is a stopping time with respect to $\mathcal{G}_t$ and $\zeta < \infty$ almost surely (a.s). A process is strong Markov with respect to $\mathcal{G} = \{\mathcal{G}\}_t$ if equation (6.4) holds for all ζ stopping times with respect to $\{\mathcal{G}_t\}$. In other words what a strong Markov process has going for it is that the Markov property not only holds at any times but also at any stopping time. This is important since it allows us to use Markov processes in very interesting arguments. Clearly any strong Markov process is Markov.

Remark 6.2.1. *Any Markov process is a strong Markov at ζ if ζ is discrete valued.*

Example 6.2.1. *The 1-D Brownian motion is a Markov process, with the transition density function:*

$$\mathcal{P}(t,x,\Gamma) = \int_\Gamma \frac{1}{\sqrt{2\pi t}} e^{-\frac{(y-x)^2}{2t}} dy.$$

The Brownian motion is one of the few examples where we can actually write the transition distribution explicitly. In general, it is not possible to calculate these transition probabilities and directly specifying them does little to the characteristic processes as well. To be able to better characterize these processes we define an operator $T(t)f(x) = \int f(y)\mathcal{P}(t, x, dy)$ and then using the Chapman-Kolmogorov equation, we obtain:

$$T(s + t)f(x) = T(t) \cdot T(s)f(x). \tag{6.5}$$

This particular expression tells us that $T(t)$ is a semigroup operator specifically a contraction operator $||T(t)|| \leq 1$.

We say that $\{X_t\}$ a Markov process corresponds to a semigroup operator $T(t)$ iff:

$$T(t)f(X_s) = \int f(y)\mathcal{P}(t, X_s, dy) = E[X_{t+s}|\mathcal{F}_s^x] \quad \forall \quad f \quad \text{measurable.}$$

This $T(t)$ and the initial distribution of X_0 completely defines the process.

In the next section, we briefly review some elements of semigroup theory.

6.3 Basic elements of semigroup theory

Let L be a Banach space of functions. A Banach space is a normed linear space that is a complete metric space with respect to the metric derived from its norm. In essence, we shall use bounded continuous functions on E. A one parameter family $\{T(t)\}_{t\geq 0}$ of bounded linear operators is called a semigroup if :

1. $T(0) = I$ (identity operator)
2. $T(s + t) = T(s) \cdot T(t) \ \forall \ t, s \geq 0$

A semigroup is called strongly continuous if:

$$\lim_{t\to 0} T(t)f = f \text{ for all } f \in L \quad (||T(t)f - f|| \to 0 \quad as \quad t \to 0) \tag{6.6}$$

A semigroup is called a contraction if $||T(t)|| \leq 1$ for all t.

Example 6.3.1. *Let $E = \mathbb{R}^n$ and let B a $n \times n$ matrix.*

Define:

$$e^{tB} = \sum_{k=0}^{\infty} \frac{1}{k!} t^k B^k,$$

where B^k are powers of the matrix which are well defined since the matrix is square and the convention $B^0 = I_n$ the identity matrix.

Then, the operator: $T(t) = e^{tB}$ forms a strongly continuous semigroup and

$$||e^{tB}|| \leq \sum_{k=0}^{\infty} \frac{1}{k!} t^k ||B^k|| \leq \sum \frac{1}{k!} t^k ||B||^k \leq e^{t||B||},$$

where the norm is the supremum norm.

In general for a strongly continuous semigroup we have $||T(t)|| \leq Me^{mt}$, $t \geq 0$ for some constants $M \geq 1$ and $m \geq 0$.

6.3.1 Infinitesimal operator of semigroup

An *Infinitesimal operator of semigroup* $T(t)$ on L is a linear operator defined as:

$$Af = \lim_{t \to 0} \frac{1}{t}(T(t)f - f).$$

The domain of definition is those functions for which the operator exists (i.e., $D(A) = \{f | Af \quad \text{exists}\}$). The infinitesimal generator acn be thought in some sense as the right derivative of the function: $t \to T(t)f$.

Theorem 6.3.1. *If A is the infinitesimal generator of $T(t)$ and $T(t)$ is strongly continuous then $\int_0^t T(s)fds \in D(A)$ and*

1. *If $f \in L$ and $t \geq 0$ then $T(t)f - f = A\int_0^t T(s)fds$*
2. *If $f \in D(A)$, then $\frac{d}{dt}T(t)f = AT(t)f = T(t)Af$ where $T(t)f \in D(A)$*
3. *$f \in D(A)$, then $T(t)f - f = \int_0^t AT(s)fds = \int_0^t T(s)Afds$*

6.3.2 Feller semigroup

A Feller semigroup is a strongly continuous, positive, contraction semigroup

$$T(t)f(x) = \int f(y)\mathcal{P}(t, x, dy) \quad \forall \quad t, x$$

This $T(t)$ corresponds to a special case of Markov process called a Feller process. More specifically,

$$E[f(X_{t+s})|\mathcal{F}_s^x] = T(t)f(X_s) = \int f(y)\mathcal{P}(t, X_s, dy).$$

Theorem 6.3.2. *Let X be a Markov process with generator A.*

1. *If $f \in D(A)$ then $M_t = f(X_t) - f(X_0) - \int_0^t Af(X_0)ds$ is a martingale.*
2. *If $f \in C_0(E)$ and there exists a function $g \in C_0(E)$ such that $f(X_t) - f(X_0) - \int_0^t g(X_s)ds$ is a martingale then $f \in D(A)$ and $Af = g$.*

Proof. In this proof for simplicity of notations we neglect $f(X_0)$. We have:

$$\begin{aligned}
E(M_{t+s}|\mathcal{F}_t) &= E[f(X_{t+s})|\mathcal{F}_t] - E\left[\int_0^{t+s} Af(X_u)du|\mathcal{F}_t\right] \\
&= T(s)f(X_t) - \int_0^t Af(X_u)du - \int_t^{t+s} E[Af(X_u)|\mathcal{F}_t]du \\
&= T(s)f(X_t) - \int_0^t Af(X_u)du - \int_0^s T(z)Af(X_t)dz.
\end{aligned}$$

But A is an infinitesimal generator and so,

$$T(t)f - f = \int_0^t T(s)Afds,$$

which implies that

$$f = T(t)f - \int_0^t T(s)Afds$$

Substituting we obtain,

$$E(M_{t+s}|\mathcal{F}_t) = T(s)f(X_t) - \int_0^s T(z)Af(X_t)dz - \int_0^t Af(X_u)du = M_t.$$

where $f(X_t) = T(s)f(X_t) - \int_0^s T(z)Af(X_t)dz$. Note there is an $f(X_0)$ missing from the expression, that is the one we neglected.

We finally set

$$T(t)f - f = \int_0^t T(s)gds,$$

which is the property of the infinitesimal generator. □

6.4 General Diffusion process

In this section, we briefly describe the general diffusion process and present some useful examples.

Let $a = (a_{ij})_{ij}$ be a continuous symmetric, nonnegative definite $d \times d$ matrix valued function on $\mathbb{R}^d$. Let $b : \mathbb{R}^d \to \mathbb{R}^d$ be a continuous function. Define:

$$Lf = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}\frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j}f + \sum_{i=1}^{d} b_i \frac{\partial}{\partial x_i}f, \text{ for all } f \in C_c^{\infty}(\mathbb{R}^d).$$

In general, for non-homogeneous process we need to define:

$$L_t f(x) = \frac{1}{2}\sum_{i,j=1} a_{ij}(t,x)\partial_{x_i}\partial_{x_j}f(x) + \sum_{i=1}^{d} b_i(t,x)\partial_{x_i}f(x).$$

A Markov process is a diffusion process with infinitesimal generator L if it has continuous paths and

$$E[f(X_t) - f(X_0)] = E\left[\int_0^t L(f(X_s))ds\right].$$

Conversely, if $\{X_t\}$ has continuous paths its generator is given by:

$$Af = c(x)f + \sum_i b_i(x)\frac{\partial f}{\partial x_i} + \sum_{ij} a_{ij}(x)\frac{\partial^2 f}{\partial x_i \partial x_j}.$$

Given the connection above, we use Itô's lemma to give a general characterization for diffusion processes.

Definition 6.4.1. If X_t solves a $d-$dimensional stochastic differential equation:

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t$$

where b is a vector and σ is a $d \times d$ matrix $(\sigma_{ij})_{i,j}$ then X_t is a diffusion (Feller) process with infinitesimal generator:

$$L_t = \sum_{j=1}^{d} b_j(X_t)\frac{\partial}{\partial x_j} + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{k=1}^{d}\sigma_{ik}(X_t)\sigma_{kj}(X_t)\frac{\partial^2}{\partial x_i \partial x_j}.$$

Note that writing $a_{ij} = \sum_{k=1}^{d} \sigma_{ik}\sigma_{kj}$, puts the process in the classical form.

We recall from Theorem 6.3.2 that for all f, $f(X_t) - f(X_0) - \int_0^t L_t f(X_t)$ is a martingale.

Next, we present some useful examples of diffusion processes.

Example 6.4.1 (Pure jump Markov process). *Let X_t be a compound Poisson process, with λ the jump intensity function and μ the probability distribution of jumps. That is in the interval $[0, t]$ the process jumps N times where N is a Poisson random variable with mean λt. Each times it jumps it does so with a magnitude Y with distribution μ. Each jump is independent. The process X_t cumulates (compounds) all these jumps. Mathematically:*

$$X_t = \sum_{i=1}^{N} Y_i.$$

This process is a Markov process and its infinitesimal generator is:

$$Af(x) = \lambda(x)\int (f(y) - f(x))\mu(x, dy)$$

Example 6.4.2 (Lévy process). *Levy processes will be discussed in details in chapter 12 of this book. Their infinitesimal generator has the form:*

$$\begin{aligned} L_t f = &\frac{1}{2}\sum_{ij} a_{ij}(t,x)\partial_i\partial_j f(x) + \sum_i b_i(t,x)\partial_i f(x) \\ &+ \int_{\mathbb{R}^d}\left(f(x+y) - f(x) - \frac{y \cdot \bigtriangledown f(x)}{1+|y|^2}\right)\mu(t, x; dy) \end{aligned}$$

In the case when jumps are not time dependent, or dependent on the current state of the system (i.e., μ is independent of (t, x)) then the integral is

$$\int_{\mathbb{R}}\left(f(x+y) - f(x) - \frac{y \cdot \bigtriangledown f(x)}{1+|y|^2}\right)\mu(dy)$$

A more specific example is presented below.

Example 6.4.3. *Suppose a stock process can be written as:*

$$dS_t = rS_t dt + \sigma S_t dW_t + S_t dJ_t,$$

where J_t is a compound Poisson process. This is a jump diffusion process.

Let

$$X_t = \log S_t$$

then using Itô lemma, under the equivalent martingale measure:

$$dX_t = \mu dt + \sigma dB_t + dJ_t,$$

where μ is $r - \frac{\sigma^2}{2} + \lambda(1 - E(e^Z))$, Z is the random variable denoting the jump rate. The process

$$J_t = \sum_{i}^{N_t} Z_i,$$

where $N_t \sim$ Poisson(λt) and Z_i are the jump rates. In this case we can write the infinitesimal generator:

$$A_t f(x) = \frac{1}{2}\sigma^2 \frac{\partial^2 f}{\partial x^2} + \mu \frac{\partial f}{\partial x} + \lambda \int_{\mathbb{R}} \left[f(x+z) - f(x)\right] p(z) dz,$$

where $p(z)$ is the distribution of jumps Z_i.

Below are some specific examples of the jump distributions $p(z)$.

Merton:

$$p(z) = \frac{1}{\sqrt{2\pi s^2}} e^{-\dfrac{(z-\mu)^2}{2s^2}}$$

and

$$E(e^z) = e^{\mu + \frac{\sigma^2}{2}}$$

Kou:

$$p(z) = p y_1 e^{-y_1 z} 1_{\{z \geq 0\}} + (1-p) y_2 e^{y_2 z} 1_{\{z \leq 0\}} \quad \text{(double exponential)}$$

and

$$\begin{aligned} E(e^z) &= 1 - (1-p)\frac{1}{y_2+1} + p\frac{1}{y_1+1} \\ &= (1-p)\frac{y_2}{y_2+1} + p\frac{y_1}{y_1-1} \end{aligned}$$

6.4.1 Option pricing application:

We recall from Theorem 6.3.2 that for all f, $f(X_t) - f(X_0) - \int_0^t L_t f(X_t)$ is a martingale. if the process X_t is Markov and L_t is its infinitesimal generator.

Suppose we want to price an option on X_t and further assume that its value is a function of the current value X_t only. This has been shown to happen for European type option when the final payoff is of the form $\psi(X_T) = F(Ke^{X_T})$. In this case the option value at time t is:

$$V_t = E[e^{-r(T-t)}\psi(x_T)|\mathcal{F}_t].$$

But since f is a function of X_t the martingale property above means that the value at t is found by solving:

$$V_t + L_t V - rV = 0,$$

which is the corresponding PDE.

6.5 A general diffusion approximation method

Ion: I need to revise this In this section, we discuss how to approximate general diffusion processes with discrete processes. This discussion forms the basics of all tree approximations. We begin with a theorem.

Theorem 6.5.1. *Let (a_{ij}) be a continuous, symmetric and nonnegative definite matrix ($d \times d$ dimensional) and $b : \mathbb{R}^d \to \mathbb{R}^d$ continuous. Let*

$$Af = \frac{1}{2}\sum a_{ij}\partial_i\partial_j f + \sum b_i \partial_i f,$$

for all $f \in C_0^{\infty(\mathbb{R}^d)}$. Assume that the martingale problem is well posed in $C_{\mathbb{R}^d}(0,\infty)$. Now let X_n and B_n be some processes, $A_n = (A_n^{ij})$ be a symmetric matrix valued process, $A_n(t) - A_n(s)$ be a nonnegative definite matrix. Let $\mathcal{F}_t^x = \sigma(X_n(s), B_n(s), A_n(s)); s \le t$. Let $\zeta_n^r = \inf\{t : |X_n(t)| \ge r \quad or \quad |X_n(t-)| \ge r\}$. Assume $M_n = X_n - B_n$ is $\{\mathcal{F}_t^x\}$ a local martingale and $M_n^i M_n^j - A_n^{ij}$ is $\{\mathcal{F}_t^x\}$ a local martingale, then $\forall r, T > 0$,

1. $\lim_{x\to\infty} E[\sup_{t\le T\wedge\zeta_n^r} |X_n(t) - X_n(t-)|^2] = 0$
2. $\lim_{x\to\infty} E[\sup |B_n(t) - B_n(t-)|^2] = 0$
3. $\lim_{x\to\infty} E[\sup |A_n^{ij}(t) - A_n^{ij}(t-)|^2] = 0$
4. $\sup_{t\le T\wedge\zeta_n^r} |B_n^i(t) - \int_0^t b_i(X_n(s))ds| \to 0$
5. $\sup |A_n^{ij}(t) - \int_0^t a_{ij}(X_n(s))ds| \to 0$

If $P \circ X_n^{-1}(0)$ implies $D \in$ probability on $\mathbb{R}^d$ (convergence in distribution) then X_n implies a solution of the martingale problem for (A, D).

Corollary 6.5.1. *Let a, b and A be defined as in theorem 6.5.1. Suppose that the martingale problem has a unique solution for each $D \in D(\mathbb{R}^d)$, let $\mu_n(x, \Gamma), \mu = 1, 2, \ldots$ be a transition function on $\mathbb{R}^d$. Let*

$$b_n(x) = n \int_{|y-x| \leq 1} (y - x) \mu_n(x, dy)$$
$$a_n(x) = n \int_{|y-x| \leq 1} (y - x)(y - x)^T \mu_n(x, dy)$$

suppose that

$$\sup_{|x| \leq r} |a_n(x) - a(x)| \to 0$$
$$\sup_{|x| \leq r} |b_n(x) - b(x)| \to 0$$
$$\sup_{|x| \leq r} n\mu(x, \{y : |y - x| \geq \epsilon\}) \to 0$$

$\forall r, \epsilon > 0$. If Y_n is a Markov chain with transition function $\mu_n(x, \Gamma)$ and define $X_n = Y_n([nt])$. If $P \circ Y_n^{-1}(0)$implies D then X_n implies the solution of the martingale problem for (A, D).

Let us consider a case where we have:

$$dX_t = \left(r - \frac{\sigma^2(Y_t)}{2}\right) dt + \sigma(Y_t) dW_t$$

$$dY_t = \alpha(m - Y_t)dt + \psi(Y_t)dz_t, f \in C_0^\infty(\mathbb{R})$$

$$b(x, y) = \begin{pmatrix} r - \frac{\sigma^2(y)}{2} \\ \alpha(m - y) \end{pmatrix}$$

$$\sigma(x, y) = \begin{pmatrix} \sigma(y) & 0 \\ 0 & \psi(y) \end{pmatrix} \Longrightarrow a(x, y) = \sigma\sigma^T = \begin{pmatrix} \sigma^2(y) & 0 \\ 0 & \psi^2(y) \end{pmatrix}$$

then

$$L_t f = \frac{1}{2}\left(\sigma^2(y)\frac{\partial^2 f}{\partial x^2} + \psi^2(y)\frac{\partial^2 f}{\partial y^2}\right) + \left(r - \frac{\sigma^2(y)}{2}\right)\frac{\partial f}{\partial x} + \alpha(m - y)\frac{\partial f}{\partial y}$$

Next consider the following SDE,

$$dS_t = \mu S_t dt + \sigma(Y_t) S_t dW_t \tag{6.7}$$

$$dY_t = \alpha(D - Y_t)dt + \psi(Y_t)dZ_t \tag{6.8}$$

Suppose $X_t = \log S_t$, then from (6.7) and (6.8), we have:

$$dX_t = \left(r - \frac{\sigma^2(Y_t)}{2}\right) dt + \sigma(Y_t) dW_t \tag{6.9}$$

$$dY_t = \alpha(D - Y_t)dt + \psi(Y_t)dZ_t \tag{6.10}$$

Assuming we know the distribution of the random variable Y_t henceforth Y, we can approximate $dX_t = (r - \frac{\sigma^2(Y)}{2})dt + \sigma(Y)dW_t$.

Let T denote an option at maturity and $\Delta t = \frac{T}{n} = h$ then we can construct a discrete Markov chain $(x(ih), \delta_{ih})$ with $P_x^z = \mathcal{P}(X_{(i+1)h} = z | X_{ih} = x)$ (depend on h).

Now define the probability measure $\mathcal{P}_x^h$ as follows:

1. $\mathcal{P}_x^h(x(0) = x) = 1$
2. $\mathcal{P}_x^h(x(t)) = \frac{(i+1)h-t}{n} x(ih) + \frac{t-ih}{n} x((i+1)h) = 1 \quad \forall \quad ih \leq t < (i+1)h$
3. $\mathcal{P}_x^h(x((i+1)h) = z | \mathcal{F}_{ih} = \mathcal{P}_x^z, \quad \forall \quad z \in \mathbb{R}, i \geq 0.$

Now:

$$b_h(x,y) = \frac{1}{h} \sum \mathcal{P}_x^z(z - x) = \frac{1}{h} E^Y[\Delta x(ih)],$$

$$a_h(x,y) = \frac{1}{h} \sum \mathcal{P}_x^z(z - x)^2 = \frac{1}{h} E^Y[(\Delta x(ih))^2].$$

Similarly:

$$b(x,y) = r - \frac{\sigma^2(y)}{2}$$

and

$$a(x,y) = \sigma^2(y).$$

With the concepts introduced above, we now have the following results.

Theorem 6.5.2. *If $b_h(x,y) \to b(x,y)$, $a_h(x,y) \to a(x,y)$ and then* $\max |z-x| = 0$.

In more general form, if $\Pi_h(x, \Gamma)$ is a transition function and we replace the probability measure 3 in the definition above with $\mathcal{P}_x^h(x(i+1)h) \in \Gamma | \mathcal{F}_{ih}) = \Pi_h(x(ih), \Gamma)$ then the theorem still works.

Remark 6.5.1. *If $A_h f(x) = \int (f(y) - f(x)) \Pi_h(x, dy)$ then $f(x(ih)) - \sum A_h f(x(jh))$ is a martingale.*

$$a_h^{ij} = \frac{1}{h} \int_{|y-x| \leq 1} (y_i - x_i)(y_j - x_j) \Pi_h(x, dy),$$

$$b_h^i = \frac{1}{h} \int_{|y-x| \leq 1} (y_i - x_i) \Pi_h(x, dy)$$

and

$$\Delta_h^\epsilon = \frac{1}{h} \Pi_h(x, \mathbb{R}^d - B(x, \epsilon))$$

assuming

$$\lim_{h \to 0} \sup_{||x|| \leq R} ||a_h(x) - a(x)|| = 0,$$

$$\lim_{h\to 0} \sup_{||x||\le R} ||b_h(x) - b(x)|| = 0,$$

and

$$\lim_{h\to 0} \sup_{||x||\le R} \Delta_h^{\epsilon}(x) = 0, \quad \forall \epsilon > 0.$$

Ion: review until here. Continuing next

6.6 Tree Methods: The Binomial tree

The binomial tree was developed during the late 1970's by [46] and perfected in the early 1980's. The binomial tree model generates a pricing tree in which every node represents the price of an underlying financial instrument at a given point in time. The methodology can be used to price options with non standard features such as path dependence and barrier events. The main idea behind the binomial tree is to construct a discrete version of the geometric Brownian motion model for the stock process S_t in (6.1).

The process in (6.1) is a continuous process. Furthermore, since it is stochastic every time one constructs a path, the path would look different. The present value of a stock price is known to be S_0. The binomial tree (and indeed any tree in fact) may be thought of forming a collection of "most probably paths". When the number of steps in the tree approximation increases the tree contains more and better paths until at the limit we obtain the same realizations as the process in (6.1).

6.6.1 One step binomial tree

We start the construction by looking at a single step. The idea is that, if this step is done properly, the extension to any number of steps is straightforward.

In order to proceed, we state the Girsanov's theorem [152].

Theorem 6.6.1. (Girsanov, One-dimensional Case) *Let $B(t), 0 \le t \le T$, be a Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $\mathcal{F}(t), 0 \le t \le T$, be the accompanying filtration, and let $\theta(t), 0 \le t \le T$, be a process adapted to this filtration. For $0 \le t \le T$, define*

$$\tilde{B}(t) = \int_0^t \theta(u)du + B(t),$$

$$Z(t) = \exp\{-\int_0^t \theta(u)dB(u) - \frac{1}{2}\int_0^t \theta^2(u)du\},$$

and define a new probability measure by

$$\tilde{\mathbb{P}}(A) = \int_A Z(T)d\mathbb{P}, \quad \forall A \in \mathcal{F}.$$

Under $\tilde{\mathbb{P}}$, the process $\tilde{B}(t), 0 \le t \le T$, is a Brownian motion.

Girsanov's theorem allows us to go from the unknown drift parameter μ to a riskfree rate r which is common to all the assets. Applying Girsanov's theorem to (6.1) and using the Itô's formula on the process $X_t = \log S_t$, we obtain the stochastic differential equation:

$$dX_t = \left(r - \frac{\sigma^2}{2}\right) dt + \sigma dW_t. \tag{6.11}$$

We note that the stochastic equation of the log process $X_t = \log S_t$ is much simpler than the original (6.1). This translates into a construction much simpler for the X_t than for the original process S_t. For each choice of the martingale measure (probabilities in the tree) a tree for S_t is going to be perfectly equivalent with a tree in S_t and vice-versa. Specifically, suppose that we have constructed a tree for the process X_t which has $x+\Delta x_u$ and $x+\Delta x_d$ steps up and down from x respectively. Then an equivalent tree for S_t could be constructed immediately by taking the next steps from $S = e^x$ as $Su = e^{x+\Delta x_u} = Se^{\Delta x_u}$ and $Sd = e^{x+\Delta x_d} = Se^{\Delta x_u}$, and keeping the probabilities identical. Similarly the tree in X_t is constructed from an S_t tree by taking $\Delta x_u = \log u$ and similarly for the down step.

Due to the obvious way in which they are constructed the resulting tree for the return X_t is called an *additive tree*, and the tree for the stock process S_t is called a *multiplicative tree*. We shall be focusing on the additive tree in the remainder of this chapter simply because the construction is simpler.

To have two successors for every node, at any point in the tree we can go up to $x+\Delta x_u$ with probability p_u or down to $x+\Delta x_d$ with probability $p_d = 1-p_u$. The one-step tree for this process is presented in Figure 6.1.

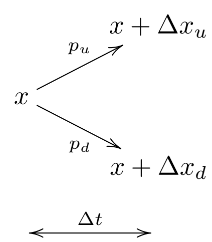

Figure 6.1: One step binomial tree for the return process

To determine the appropriate values for Δx_u, Δx_d, p_u, and p_d we use the diffusion approximation presented in section 6.5 for the model in (6.11). Recall that means the infinitesimal generator for the discrete process needs to converge to the infinitesimal generator of the continuous process. However, the functions b and a are just the mean increase and variance increase of the process. So we impose that the increase (Δx) from the tree and the model should have first two moments equal.

This is a sufficient condition since the process X_t in (6.11) is a normal random variable for any moment t fixed. We also know that a normal variable is entirely characterized by its mean and variance. Thus, equating the mean and variance increase for an infinitesimal time step will be sufficient to make sure that the binomial process will converge to the path given by the process (6.11). Specifically, the conditions we need to enforce are:

$$\begin{cases} p_u \Delta x_u + p_d \Delta x_d &= \left(r - \frac{\sigma^2}{2}\right) \Delta t \\ p_u \Delta x_u^2 + p_d \Delta x_d^2 &= \sigma^2 \Delta t + \left(r - \frac{\sigma^2}{2}\right)^2 \Delta t^2 \\ p_u + p_d &= 1 \end{cases} \tag{6.12}$$

The system (6.12) has 3 equations and 4 unknowns, thus it has an infinite number of solutions. Indeed the binomial tree approximation is not unique. In practice we have a choice of a parameter - that choice will create a specific tree.

We note that for any choice of Δx_u and Δx_d the tree recombines. For example first step successors are $x + \Delta x_u$ and Δx_d. At the second step their successors will be:

$$x + 2\Delta x_u \text{ and } x + \Delta x_u + \Delta x_d \text{ for the upper node,}$$
$$\text{and } x + \Delta x_d + \Delta x_u \text{ and } x + 2\Delta x_d \text{ for the lower node.}$$

Clearly one of the nodes is identical and the tree recombines. The recombining feature in a tree is crucial. For example, consider a binomial tree with $n = 10$ steps that recombines and one that does not recombine. The recombining tree has a total number of nodes: $1+2+3+\cdots+11 = 11\times 10/2 = 55$ nodes, whereas the one that does not recombine has $1+2+2^2+2^3+\cdots+2^{10} = 2^{11}-1 = 2047$ nodes. The latter has a much bigger number of calculations that quickly become unmanageable.

As long as the resulting probabilities in (6.12) are positive the corresponding trees constructed with the particular choice of parameters are appropriate. Some popular choices are obtained by taking $\Delta x_u = \Delta x_d$ in the system (6.12). This produces what is called the Trigeorgis tree (see [193]) which supposedly has better approximation power.

Solving the system (6.12) by taking $\Delta x_u = \Delta x_d$ will yield:

$$\begin{cases} \Delta x &= \sqrt{\left(r - \frac{\sigma^2}{2}\right)^2 \Delta t^2 + \sigma^2 \Delta t} \\ p_u &= \frac{1}{2} + \frac{1}{2} \frac{\left(r - \frac{\sigma^2}{2}\right)\Delta t}{\Delta x} \end{cases} \tag{6.13}$$

The corresponding multiplicative tree for the S_t process obtained by setting the nodes as $S_t = \exp(X_t)$ and keeping all probabilities the same is the famous Cox-Ross-Rubinstein (CRR) tree (see [46]).

Another popular tree is obtained by setting the probabilities of jumps in (6.12) equal (i.e., $p_u = p_d = 1/2$). In this case the tree for S is called Jarrow Rudd tree (see [109]).

Remarks

- The system (6.12) could be further reduced by subtracting the constant term $\left(r - \frac{\sigma^2}{2}\right) \Delta t$ from each node, the resulting tree will be much simpler.

- Finding the probabilities, p_u and p_d is equivalent to finding the martingale measure under which the pricing is made. For each solution one can construct an additive tree (moving from x to $x + \Delta x$ and $x - \Delta x$ with probabilities p_u and p_d respectively) or a multiplicative tree (moving from S to Su and Sd with probabilities p_u and p_d respectively). The two constructed trees are completely equivalent.

To construct a full tree we need to first solve the one step tree for a time interval $\Delta t = T/n$ where T is the maturity of the option to be priced and n is a parameter of our choice. We then use this tree to price options as in the next sections.

6.6.2 Using the tree to price an European Option

After constructing the tree, we basically have possible paths of the process X_t or equivalently the process $S_t = \exp(X_t)$. Since these are possible paths we can use the binomial tree to calculate the present price of any path dependent option.

We begin this study by looking at the European Call. The European Call is written on a particular stock with current price S_0, and is characterized by maturity T and strike price K.

The first step is to divide the interval $[0, T]$ into $n+1$ equally spaced points, then we will construct our tree with n steps, $\Delta t = T/n$ in this case. The times in our tree will be: $t_0 = 0, t_1 = T/n, t_2 = 2T/n, \ldots, t_n = nT/n = T$.

Next we construct the return tree (X_t) as above starting with $X_0 = x_0 = \log(S_0)$, and the last branches of the tree ending in possible values for $X_{t_n} = X_T$. We remark that since we constructed the possible values at time T we can calculate for every terminal node in the tree the value of the Call option at that terminal node using:

$$C(T) = (S_T - K)_+ = \left(e^{X_T} - K\right)_+ \tag{6.14}$$

Thus now we know the possible payoffs of the option at time T. Suppose we want to calculate the value of the option at time $t = 0$. Using the Girsanov theorem and the Harrison and Pliska result [85], the discounted price of the process $\{e^{-rt}C(t)\}$ is a continuous time martingale. Therefore we may write that the value at time 0 must be:

$$C(0) = \mathbf{E}\left[e^{-rT} C(T) | \mathscr{F}_0\right].$$

Now we can use the basic properties of conditional expectation and the fact that

$T/n = \Delta t$ or $e^{-rT} = \left(e^{-r\Delta T}\right)^n := \delta^n$ to write:

$$\begin{aligned} C(0) &= \mathbf{E}\left[\delta^n C(T) | \mathscr{F}_0\right] = \mathbf{E}\left[\delta^{n-1} \mathbf{E}\left[\delta C(T) | \mathscr{F}_1\right] | \mathscr{F}_0\right] \\ &= \delta \mathbf{E}\left[\delta \mathbf{E}\left[\ldots \delta \mathbf{E}\left[\delta C(T) | \mathscr{F}_{n-1}\right] | \ldots \mathscr{F}_1\right] | \mathscr{F}_0\right], \end{aligned}$$

where $\delta = e^{-r\Delta t}$ is a discount factor.

This formula allows us to recursively go back in the tree toward the time $t = 0$. When this happens, we will eventually reach the first node of the tree at C_0 and that will give us the value of the option. More precisely since we know the probabilities of the up and down steps to be p_u and p_d respectively, and the possible Call values one step ahead to be C^1 and C^2, we can calculate the value of the option at the previous step as:

$$C = e^{-r\Delta t}\left(p_d C^1 + p_u C^2\right). \tag{6.15}$$

What is remarkable about the above construction is that we constructed the return tree just to calculate the value of the option at the final nodes. Note that when we stepped backwards in the tree we did not use the intermediate values of the stock, we only used the probabilities. This is due to the fact that we are pricing the European option which only depends on the final stock price. This situation will change when pricing any path dependent option such as the American, Asian, barrier, etc.

Furthermore, we do not actually need to go through all the values in the tree for this option. For instance, the uppermost value in the final step has probability p_u^n. The next one can only be reached by paths with probability $p_u^{n-1} p_d$ and there are $\binom{n}{1}$ of them. Therefore, we can actually write the final value of the European call as:

$$C = \sum_{i=0}^{n} \binom{n}{i} p_u^{n-i} p_d^i (e^{x+(n-i)\Delta x_u + i\Delta x_d} - K)_+$$

as an alternative to going through the tree.

6.6.3 Using the tree to price an American Option

The American option can be exercised at any time, thus we will need to calculate its value at $t = 0$ and the optimal exercise time τ. Estimating τ is not a simple task, but in principle it can be done. τ is a random variable and therefore it has an expected value and a variance. The expected value can be calculated by looking at all the points in the tree where it is optimal to early exercise the option and the probabilities for all such points. Then we can calculate the expected value and variance.

For the value of the American option we will proceed in similar fashion as for the European option. We will construct the tree in the same way we did for the European option and then we will calculate the value of the option at the terminal nodes. For example for an American Put option we have:

$$P(T) = (K - S_T)_+ = \left(K - e^{X_T}\right)_+ \tag{6.16}$$

Then we recursively go back in the tree in a similar way as we did in the case of the European option. The only difference is that for the American option, we have the early exercise condition. So at every node we have to decide if it would have been optimal to exercise the option rather than to hold onto it. More precisely using the same notation as before, we again calculate the value of the option today if we would hold onto it as:

$$C = e^{-r\Delta t}\left(p_d C^1 + p_u C^2\right).$$

But what if we actually exercised the option at that point in time, say t_i? Then we would obtain $(K - S_{t_i})_+$. Since we can only do one of the two, we will obviously do the best thing, so the value of the option at this particular node will be the maximum of the two values. Then again, we recursively work the tree backward all the way to $C(0)$ and that will yield the value of the American Put option. A possible value of the optimal exercise time τ is encountered whenever the value obtained by exercising early $(K - S_{t_i})_+$ is larger than the expectation coming down from the tree.

We now discuss a more general approach of pricing path dependent options using the binomial tree techniques.

6.6.4 Using the tree to price any path dependent option

From the previous subsection we see that the binomial tree could be applied to any path dependent option. The trick is to keep track of the value of the option across various paths of the tree.

Please see [38] for a detailed example of a down and out American option. This is an American option (like before it can be exercised anytime during its lifespan). In addition if the asset price on which the option is dependent falls below a certain level H (down) then the option is worthless, it ceases to exist (out).

Binomial trees for dividend paying asset are very useful, however they are not sufficiently different from the regular binomial tree construction. The model developed thus far can be modified easily to price options on underlying assets other than dividend and non-dividend-paying assets.

Other applications of the binomial tree methodology is in the computation of hedge sensitivities. A brief introduction is presented below. Please refer to [38] for more details. In this application, we discuss how the binomial tree can be used to compute hedge sensitivities.

6.6.5 Using the tree for computing hedge sensitivities - the Greeks

When hedging an option as we discussed earlier in chapter 3 section 3.3, it is important to calculate the changes in option price with the various parameters in the model. We also recall in 3 section 3.3 that these derivatives are denoted with *delta* Δ, *gamma* Γ, *Vega, theta* Θ and *rho* ρ.

Their mathematical definition is given as:

$$\begin{cases} \Delta = \frac{\partial C}{\partial S} \\ \Gamma = \frac{\partial^2 C}{\partial S^2} \\ \text{Vega} = \frac{\partial C}{\partial \sigma} \\ \rho = \frac{\partial C}{\partial r} \\ \Theta = \frac{\partial C}{\partial t} \end{cases} \tag{6.17}$$

In order to approximate the greeks, one can use the method discussed by [38] that uses only one constructed binomial tree to calculate an approximation for Δ and Γ one step in the future.

Alternatively, one can construct two binomial trees in the case of Δ and three in the case of Γ starting from slightly different initial asset prices and use the same formula to approximate Δ and Γ at the present time.

In order to approximate Θ derivative, it requires knowledge of stock prices at two moments in time. It could be calculated using a similar approach as for $Vega$ and ρ but since option price varies in time with the Brownian motion driving the process S_t, it is unnecessary to calculate the slope of the path.

6.6.6 Further discussion on the American option pricing

We consider a multiplicative binomial tree:

$$\begin{array}{cccc} & & Suu & ... \\ & Su & & \\ & & & ... \\ S & & Sud & \\ & & & \\ & Sd & & \\ & & Sdd & ... \end{array}$$

In the following examples we work with a probability of an up step:

$$\frac{Se^{r\Delta t} - S_d}{S_u - S_d} = \frac{e^{r\Delta t} - d}{u - d} = p$$

Example: European Call option.
Suppose a strike price $K = 105$, $u = 1$, $d = 0.9$ and $r = 7\%$ (annual) with T semesters and $\Delta T = \frac{1}{2}$. Suppose that the original price of the asset is $S = 100$.

In this case, we have the following tree:

$$
\begin{array}{ccc}
 & & 121 \\
 & & (16) \\
 & 110 & \\
 & 10,4764 & \\
100 & & 99 \\
(6,8597) & & (0) \\
 & 90 & \\
 & (0) & \\
 & & 81 \\
 & & (0)
\end{array}
$$

The probability is then $p = 0.6781$. Next, when we consider an European Put option, the situation is different. In this case, we have:

$$
\begin{array}{ccc}
 & & 121 \\
 & & (0) \\
 & 110 & \\
 & (1,865) & \\
100 & & 99 \\
(4,772) & & (6) \\
 & 90 & \\
 & (11,3885) & \\
 & & 81 \\
 & & (24)
\end{array}
$$

Remark 6.6.1. *We observe that there exists a relation between the values of the Put and the Call options:*

$$C_t + Ke^{-r(T-t)} = P_t + S_t.$$

This relation is called the Put-Call parity *as discussed in chapter 3 section 3.9.*

We know that it is possible to exercise an American Call option at any time or, in the discrete case, at any node. We compute the maximum between 0 and $K - S_t$ i.e.

$$P_t = \max\{0, K - S_t\}.$$

We recall that in a Call option the relation is the reverse of the previous relation i.e.:

$$C_t = \max\{0, S_t - K\}.$$

We recall in chapter 2 that without considering dividends, the values of an American and European Call options are the same (i.e. $C_E = C_A$). Suppose that we exercise at time t, (i.e.: $S_t - K > 0$) and at a future time T, it grows with a risk-free rate $e^{r(T-t)}$. As the expected payoff is $E(S(T)) = S_t e^{r(T-t)} > S_t e^{r(T-t)} - K$, then it is not convenient to exercise, because the value will go up. We recall that in chapter 3 section 3.9, we can write the Put-Call parity as:

$$C_t - P_t = S_t - Ke^{-r(T-t)}.$$

For $t = T$,

$$\begin{aligned} C_T &= \max\{S_T - K, 0\} \\ P_T &= \max\{K - S_T, 0\}, \end{aligned}$$

so that we have the equality:

$$C_T - P_T = S_T - K$$

Now consider the relation:

$$\pi_t = \begin{cases} 1 \text{ Call} & \text{in long} \\ 1 \text{ Put} & \text{in short} \\ 1 \text{ share} & \text{in short} \end{cases}$$

Then we have that,

$$\begin{aligned} \pi_t &= C_t - P_t - S_t \\ \pi_T &= C_T - P_T = S_T - K - S_T = -K \end{aligned}$$

and assuming no arbitrage,

$$\pi_t = e^{-r(T-t)} \pi_T.$$

In general, suppose that at time T, we have $n+1$ possible states i.e. $T = n\Delta t$. We assume that,

$$S^j(T) = Su^j d^{n-j}$$

with

$$p = \frac{e^{r\Delta t} - d}{u - d}.$$

Then the probability that S is in the state S^j at time T is given as

$$P\left(S(T) = S^j(T)\right) = \begin{pmatrix} n \\ j \end{pmatrix} p^j (1-p)^{n-j}.$$

and using the Newton binomial formula, the expected value is:

$$\begin{aligned} E(S(T)) &= \sum_{j=0}^{n} \begin{pmatrix} n \\ j \end{pmatrix} p^j (1-p)^{n-j} Su^j d^{n-j} \\ &= S(pu + (1-p)d)^n. \end{aligned}$$

On the other hand,

$$pu + (1-p)\,d = p\,(u-d) + d = e^{r\Delta t}$$

and from $T = n\Delta t$ we obtain

$$E\,(S\,(T)) = S\,(pu + (1-p)\,d)^n = Se^{rn\Delta t} = Se^{rT}.$$

This implies that the expected value is proportional to the basis value and the proportion depends on the interest rate.

Now suppose we have the same situation with an European Call with strike price K and the relation:

$$C_j = C_{u^j d^{n-j}} = \max\{Su^j d^{n-j} - K, 0\},$$

then

$$\begin{aligned} E\,(C\,(T)) &= \sum_{j=0}^{n} \binom{n}{j} p^j\,(1-p)^{n-j}\,C_j \\ &= \sum_{j=n_0}^{n} \binom{n}{j} p^j\,(1-p)^{n-j}\,C_j \end{aligned}$$

where

$$n_0 = \min\{j : C_j > 0\}$$

Then substituting C_j, in the above equation, we have

$$\begin{aligned} E\,(C\,(T)) &= \sum_{j=n_0}^{n} \binom{n}{j} p^j\,(1-p)^{n-j}\left(Su^j d^{n-j} - K\right) \\ &= S\sum_{j=n_0}^{n} \binom{n}{j} (pu)^j\,((1-p)\,d)^{n-j} - KP\,(U \geq n_0) \end{aligned}$$

where U is a stochastic variable that follows a binomial distribution with mean $E\,(U) = np$ and variance $Var\,(U) = np\,(1-p)$.

Using the central limit theorem we can approximate the binomial distribution by a normal distribution and obtain:

$$P\,(U \geq n_0) \approx P\left(Z \geq \frac{n_0 - np}{\sqrt{np\,(1-p)}}\right).$$

6.6.7 Wiener process (Brownian motion)

We begin our discussion on the Wiener process with an empirical analysis. Assuming a random walk with probability $\frac{1}{2}$, we will have two variables: the time $t \in [0,T]$ and the position $x \in [-X, X]$. Intuitively, we recall that a *Markov process* does not have memory. Thus the present is only relevant.

For each $j = 1, 2, ..., n$, consider

$$u_{k,j} = P\left(x_k = j\right),$$

where k represents the time and j the position.

We recall that for $P(B) \neq 0$ we can define the conditional probability of the event A given that the event B occurs as:

$$P\left(A|B\right) = \frac{P\left(A \cap B\right)}{P\left(B\right)}.$$

To the position j in time $k+1$ we can arrive only from the position $j-1$, or j at time k, so we have:

$$u_{k+1,j} = \frac{1}{2}\left(u_{k,j} + u_{k,j-1}\right). \tag{6.18}$$

We can rewrite (6.18) as:

$$u_{k+1,j} = \frac{1}{2}\left[\left(u_{k,j+1} - u_{k,j}\right) - \left(u_{k,j} - u_{k,j-1}\right)\right] + u_{k,j}$$

or

$$u_{k+1,j} - u_{k,j} = \frac{1}{2}\left[\left(u_{k,j+1} - u_{k,j}\right) - \left(u_{k,j} - u_{k,j-1}\right)\right].$$

Using the notation,

$$u_{k,j} = u\left(t_k, x_j\right),$$

we obtain:

$$u\left(t_{k+1}, x_j\right) - u\left(t_k, x_j\right) = \frac{1}{2}\left[\left(u\left(t_k, x_{j+1}\right) - u\left(t_k, x_j\right)\right) - \left(u\left(t_k, x_j\right) - u\left(t_k, x_{j-1}\right)\right)\right]. \tag{6.19}$$

Now let Δt and Δx be such that $X = n\Delta x, T = n\Delta t$ and then $\frac{X}{\Delta x} = \frac{T}{\Delta t}$. Then multiplying (6.19) by $\frac{1}{\Delta t}$ we obtain:

$$\frac{1}{\Delta t}\left(u\left(t_{k+1}, x_j\right) - u\left(t_k, x_j\right)\right) = \frac{1}{2\Delta t}\left[\left(u\left(t_k, x_{j+1}\right) - u\left(t_k, x_j\right)\right) - \left(u\left(t_k, x_j\right) - u\left(t_k, x_{j-1}\right)\right)\right]. \tag{6.20}$$

If we take $\Delta t \to 0$ the first term in (6.20) converges to $\partial_t\left(u\right)$. For the second term, if we assume that:

$$\Delta t \approx \left(\Delta x\right)^2$$

taking into account that $\Delta t = \frac{T\Delta x}{X}$

$$\frac{1}{2\Delta t} = \frac{X}{2T\Delta x}$$

we can conclude that the second term converges to $\partial_{xx}\left(u\right)$. So, from the random walks we get a discrete version of the heat equation

$$\partial_t\left(u\right) = \frac{1}{2}\partial_{xx}\left(u\right)$$

As an example, consider the random walk with step $\Delta x = \sqrt{\Delta t}$, that is:

$$\begin{cases} x_{j+1} = x_j \pm \sqrt{\Delta t} \\ x_0 = 0 \end{cases}$$

We claim that the expected value after n steps is zero. The stochastic variables $x_{j+1} - x_j = \pm\sqrt{\Delta t}$ are independents and so

$$E(x_n) = VE\left(\sum_{j=0}^{n-1}(x_{j+1} - x_j)\right) = \sum_{j=0}^{n-1} VE(x_{j+1} - x_j) = 0$$

and

$$Var(x_n) = \sum_{j=0}^{n-1} Var(x_{j+1} - x_j) = \sum_{j=0}^{n-1} \Delta t = n\Delta t = T.$$

If we interpolate the points $\{x_j\}_{1\leq j\leq n-1}$ we obtain

$$x(j) = \frac{x_{j+1} - x_j}{\Delta t}(t - t_j) + x_j \tag{6.21}$$

for $t_j \leq t \leq t_{j+1}$. (6.21) is a *Markovian process*, because:

1. $\forall a > 0, \{x(t_k + a) - x(t_k)\}$ is independent of the history $\{x(s) : s \leq t_k\}$
2. $$\begin{aligned} E(x(t_k + a) - x(t_k)) &= 0 \\ Var(x(t_k + a) - x(t_k)) &= a \end{aligned}$$

Remark 6.6.2. *If $\Delta t \ll 1$ then $x \approx N(0, \sqrt{a})$ (i.e normal distributed with mean 0 and variance a) and then,*

$$P(x(t + a) - x(t) \geq y) \approx \frac{1}{\sqrt{2\pi}} \int_y^\infty e^{-\frac{t^2}{2a}} dt.$$

This is due to the Central Limit Theorem that guarantees that if N is large enough, the distribution can be approximated by Gaussian.

In the next section, we will study the pricing of options when the volatility varies. We will consider models of the form,

$$dS = \mu S dt + \sigma S dZ.$$

The parameter σ will be the variance or volatility. When $\sigma = 0$ the model is deterministic. The parameters μ and σ depend on the asset and we have to estimate them. We can estimate them from the historical data for the assets.

In the next section, we relate the tree method to the Black-Scholes (B-S) model.

6.7 The Black-Scholes (B-S) model and the tree method.

We recall in chapter 3 that the derivative price for the B-S model is given as:

$$rV = Sr\frac{\partial V}{\partial S} + \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}.$$

Also in section 6.6 the discrete binomial model was modeled as:

$$\begin{array}{ccc} & & S_u \\ & & (V_u) \\ & \nearrow_p & \\ S & & \\ & \searrow_{1-p} & S_d \\ & & (V_d) \end{array}$$

with portfolio,

$$\pi = \Delta S - V$$

where Δ was chosen such that

$$\Delta = \frac{V_u - V_d}{Su - Sd}$$

and $\pi_u = \pi_d$. After taking the limit to the continuous, we obtained:

$$p = \frac{e^{rT} - d}{u - d}.$$

Solving for V, we obtain:

$$V = e^{-rT}\left(pV_u + (1-p)\,V_d\right). \tag{6.22}$$

In the B-S model we do not have only two states and so in order to eliminate the stochastic contribution, we set $\Delta = \frac{\partial V}{\partial S}$. We know that $\log\left(\frac{S_T}{S}\right)$ is normally distributed, i.e.

$$\log\left(\frac{S_T}{S}\right) \sim N\left(\left(\mu - \frac{\sigma^2}{2}\right)(T-t)\,;\sigma^2\,(T-t)\right) = N\left(M, \sigma^2\,(T-t)\right)$$

where $M = \left(\mu - \frac{\sigma^2}{2}\right)(T-t)$. Therefore, if we know the distribution we can calculate the expected value in a "fixed state" (S,t) i.e. $E_{(S,t)}\,(\text{Payoff})$. Thus from (6.22) we would have that:

$$V = e^{-r(T-t)}E_{(S,t)}\,(\text{Payoff}).$$

Remarks 6.7.1. *1. In (6.22), p is not a probability. It is just a number between zero and one and therefore we can choose $u = \frac{1}{d}$ and $p = \frac{1}{2}$.*

2. *The B-S model does not depend on* μ*. It only depends on* σ *and* r*. Therefore if* μ *is very large, the value of the asset will also be very large, and this may be a factor of distortion for the value of* V*.*

3. *How can we compare* V *with* $e^{-r(T-t)}E_{(S,t)}\,(Payoff)$*? If* $\mu \gg 0$ *then*

$$V < e^{-r(T-t)}E_{(S,t)}\,(Payoff)$$

Hence, from

$$P\left(\log\left(\frac{S_T}{S}\right) \leq \xi\right) = \frac{1}{\sqrt{2\pi}\sigma\,(T-t)}\int_{-\infty}^{\xi} e^{-\frac{(x-M)^2}{2\sigma^2(T-t)}}dx$$

we obtain the distribution of S_T *to be*

$$P\left(S_T \leq Se^{\xi}\right) = \frac{1}{\sqrt{2\pi}\sigma\sqrt{(T-t)}}\int_0^{Se^{\xi}} \frac{1}{w}e^{-\frac{\left(\log\left(\frac{w}{S}\right)-M\right)^2}{2\sigma^2(T-t)}}dw$$

so that

$$f_{S_T} = \frac{1}{\sqrt{2\pi}\sigma\sqrt{(T-t)}}\frac{1}{w}e^{-\frac{\left(\log\left(\frac{w}{S}\right)-M\right)^2}{2\sigma^2(T-t)}}$$

is the density of S_T*.*

If we consider an European Call option the payoff of S *at* T *is* $(S-K)^+$ *hence,*

$$E_{(S,t)}\,(Payoff) = \frac{1}{\sqrt{2\pi}\sigma\sqrt{(T-t)}}\int_K^{A} \frac{1}{w}e^{-\frac{\left(\log\left(\frac{w}{S}\right)-M\right)^2}{2\sigma^2(T-t)}}\,(w-K)\,dw$$

By a convenient substitution we can see that if $\mu \to +\infty$ *then* $E_{(S,t)}\,(Payoff) \to +\infty$*, therefore for* T *large enough we will have:*

$$V < e^{-r(T-t)}E_{(S,t)}\,(Payoff)$$

4. *It is reasonable to have the condition:* $\mu - \frac{\sigma^2}{2} > r$*. Suppose we construct a risk-free portfolio,*

$$d\pi = r\pi dt.$$

Because of the non-arbitrage condition, we have that for a given bond B *(risk-free asset)* $dB = rBdt$*. Recalling the Put-Call parity i.e.*

$$C - P = S - Ke^{-r(T-t)},$$

intuitively when we have a portfolio

$$\pi = \begin{cases} 1 & \textit{Call (long)} \\ 1 & \textit{Put (short)} \end{cases}$$

then,

$$\pi_T = (S_T - K)^+ - (K - S_T)^+ = S_T - K \tag{6.23}$$

Thus (6.23) is not correct since the non-arbitrage condition can be applied to a risk free asset. The correct deduction is presented as follows; if we want to have the present value of the asset, we begin by constructing a risk-free portfolio:

$$\pi = \begin{cases} 1 & \text{Call (long)} \\ 1 & \text{Put(short)} \\ 1 & \text{share (short)} \end{cases}$$

Indeed, if $\pi = C - P - S$ and $\pi_T = S_T - K - S_T = -K$ then we have that $\pi = e^{-r(T-t)}\pi_T$ and so

$$C - P = S - Ke^{-r(T-t)}$$

Remark 6.7.1. *We can only "go back on time" with risk free assets.*

In the following remark, we discuss another way to obtain the B-S equations. It is derived by constructing a portfolio that replicates an option.

Remark 6.7.2. *We say that a portfolio "Replicates an option" when it has the value of the option. Let* π *be defined as*

$$\begin{cases} \pi = \pi_1 S + \pi_2 B \\ \pi_u = V_u \\ \pi_d = V_d \end{cases}$$

so that,

$$\begin{array}{ccc} & & S_u \\ & & (V_u) \\ & \nearrow_p & \\ S & & \\ (V) & & \\ & \searrow_{1-p} & S_d \\ & & (V_d) \end{array}$$

Then from the system we have

$$\begin{aligned} \pi_1 Su + \pi_2 e^{rT} B &= V_u \\ \pi_1 Sd + \pi_2 e^{rT} B &= V_d \end{aligned}$$

Next solving for π_1 *and* π_2 *we obtain*

$$\pi_1 = \frac{V_u - V_d}{Su - Sd} = \Delta.$$

and

$$\pi_2 B = e^{-rT}\left(V_u - \frac{V_u - V_d}{u - d}u\right)$$

Thus the value for this portfolio is given as

$$\begin{aligned} \pi &= \frac{V_u - V_d}{u - d} + e^{-rT}\left(V_u - \frac{V_u - V_d}{u - d}u\right) \\ &= e^{-rT}\left(pV_u + (1 - p)V_d\right) \end{aligned}$$

where

$$p = \frac{e^{-rT} - d}{u - d}.$$

For example considering the continuous case:

$$dS = \mu S dt + \sigma S dZ,$$

we want to construct a portfolio that finances itself i.e.

$$\pi = \pi_1 S + \pi_2 B$$

and

$$d\pi = \pi_1 dS + \pi_2 dB.$$

Even if this is a risky portfolio (i.e. there are chances that combination of assets within individual group of investments fail to meet financial objectives) and we can not proceed as before, using Itô's 's lemma we get the same value. We want to find the value of $V(S,t)$ *i.e.*

$$dV = \sigma S \frac{\partial V}{\partial S} dZ + \{\frac{\partial V}{\partial t} + \mu S \frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\} dt.$$

If we want the portfolio to replicate the option, we need

$$\begin{cases} \pi = V \\ d\pi = dV \end{cases}$$

Then, we have that:

$$\begin{aligned} d\pi &= \pi_1 dS + \pi_2 dB \\ &= \pi_1 (\mu S dt + \sigma S dZ) + \pi_2 dB = dV. \end{aligned}$$

Since the stochastic components are equal, so we conclude that:

$$\pi_1 \sigma S = \sigma S \frac{\partial V}{\partial S}.$$

Thus:

$$\pi_1 = \frac{\partial V}{\partial S}.$$

On the other hand,

$$\pi_2 dB = \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right) dt.$$

We recall that if B *is a risk-free bond i.e.* $dB = rBdt$ *then,*

$$\pi_2 r B = \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}. \tag{6.24}$$

Adding $r\pi_1 S$ *to* (6.24) *and taking account of the condition that* $\pi = V$, *we obtain*

$$\begin{aligned} rV &= r\pi = r\pi_1 S + r\pi_2 B = \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + r\pi_1 S \\ &= \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + r\frac{\partial V}{\partial S} S \end{aligned}$$

which is the the B-S equation.

Until now, the B-S model has been derived in two different ways:

1. In chapter 3, it was derived by constructing a risk-free portfolio.
2. In remark 6.7.2, it was obtained by constructing a portfolio that replicates an option.

Next we revisit the example of the return process discussed earlier in chapter 2 of this book.

Define the SDE,

$$\begin{cases} dX = \mu(x,t)\,dt + \sigma(x,t)\,dZ \\ X(t) = x \end{cases} \tag{6.25}$$

where (x, t) is fixed and it is the first point of the stochastic process X. Define a variable u such that u takes the place of x in the stochastic process, so that:

$$u = u(X(t), t) \tag{6.26}$$

Applying Itô's lemma on (6.26) we obtain:

$$du = \frac{\partial u}{\partial t}dt + \frac{\partial u}{\partial x}dX + \frac{1}{2}\frac{\partial^2 u}{\partial x^2}dX^2 + R. \tag{6.27}$$

Squaring both sides of (6.25), we obtain:

$$dX^2 = \mu^2 dt^2 + 2\mu\sigma dt dZ + \sigma^2 dZ^2 \tag{6.28}$$

From (6.28), the only term that is going to remain is $\sigma^2 dZ^2$ since we ignore all lower terms. Now, using the fact that $dZ^2 \approx dt$ and replacing dX we obtain:

$$du = \frac{\partial u}{\partial x}\sigma dZ + \{\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x}\mu + \frac{1}{2}\sigma^2\frac{\partial^2 u}{\partial x^2}\}dt, \tag{6.29}$$

where u is a solution. In order to eliminate the second term on the right hand side of (6.29), we fix α and β such that, $\mu = \beta$ and $\frac{1}{2}\sigma^2 = \alpha$. This results in $\{\frac{\partial u}{\partial t} + \beta\frac{\partial u}{\partial x}\mu + \alpha\frac{\partial^2 u}{\partial x^2}\}dt$ which equals zero by (6.25). Therefore (6.29) reduces to

$$du = \frac{\partial u}{\partial x}\sigma dZ. \tag{6.30}$$

Integrating both sides of (6.30) between the interval t and T we obtain:

$$u\left(X\left(T\right),T\right)-u\left(x\left(t\right),t\right)=\int_t^T\frac{\partial u}{\partial x}\sigma dZ \tag{6.31}$$

Without loss of generality, we take $u\left(x\left(t\right),t\right)=u(x,t)$. Taking the expected value of (6.31) results in,

$$E\left(u\left(X\left(T\right),T\right)-u\left(x,t\right)\right)=0, \tag{6.32}$$

since the increments are Gaussian with mean 0. Hence we obtain the $Feyman-K\bar{a}c.$ formula (see [152]):

$$u\left(x,t\right)=E\left(u\left(X\left(T\right),T\right)\right)=E\left(u\left(\phi\left(x\right),T\right)\right).$$

The method described in the above example can also be used for elliptic equations by making the right change of variables. We can also perform numerical a implementation on the above example by:

1. Simulating several stochastic processes by using the value of ϕ in $X\left(t\right)$,
2. Computing an average.
3. Estimating the expected value.

These enables one to obtain a theoretical result that is analogous to the classic one for harmonic functions. Harmonic functions have the following mean-value property which states that the average value of the function over a ball or sphere is equal to its value at the center. The mean value property characterizes harmonic functions and has a remarkable number of consequences. For instance, harmonic functions are smooth because local averages over a ball vary smoothly as the ball moves. Please see [15] for more details.

We present some examples as follows:

1. Consider the equation,

$$\begin{cases} u''=1 \\ u\left(-1\right)=u\left(1\right)=0 \end{cases}$$

We fix $x\in(-1,1)$ and define the stochastic process Z (Brownian motion) with $Z\left(0\right)=x$, then since the stochastic process Z does not depend on t , $u''=1$ and applying Itô's lemma we obtain:

$$du=du\left(Z\left(t\right)\right)=u'dZ+\frac{1}{2}u''dt=u'dZ+\frac{1}{2}dt.$$

Next, integrating between 0 and T, we obtain:

$$u\left(Z\left(T\right)\right)-u\left(x\right)=\int_0^T u'dZ+\int_0^T\frac{1}{2}dt.$$

If we let T_x be the time so that the process arrives at a boundary then it is possible to prove that this happens with probability one. Thus, using T_x instead of T we have that $u(Z(T_x)) = 0$ and we can therefore conclude that:

$$u(x) = -\frac{1}{2} E_x(T_x). \tag{6.33}$$

As a consequence we know the expected value of the time the process arrives at the boundary, because from the solution of our differential equation, we have

$$u = -\frac{1}{2}\left(1 - x^2\right). \tag{6.34}$$

Comparing equation 6.33 to equation 6.34, we observe that

$$E_x(T_x) = 1 - x^2$$

In general, the same computation can be performed for the equation in $\mathbb{R}^n$

$$\begin{cases} \Delta u = 1 \\ u|_{\partial\Omega} = 0 \end{cases}$$

In this case, we define a Brownian motion in $\mathbb{R}^n$ with multivariate normal distribution.

2. Consider the equation,

$$\begin{cases} u'' = 0 \\ u(-1) = a; u(1) = b \end{cases}$$

Let Z be a Brownian motion so that $Z(0) = x$. In the same way as in the previous example, applying Itô's lemma to $u(Z(t))$ we obtain:

$$du = u' dZ.$$

Integrating between 0 and T results in,

$$u(Z(t)) - u(x) = \int_0^T u' dZ.$$

We recall that $Z(T) \sim N(0, T)$. Therefore the time for arriving at the boundary is T_x, hence $Z(T_x) = \pm 1$. This shows that the solution is linear.

3. Consider the equation

$$u'' = 0$$

with initial point in $x \in (x - r, x + r)$. As in the previous examples: because we are considering a Brownian motion, the probability of arriving at the boundary is $\frac{1}{2}$ and so the solution of the equation is

$$u(x) = \frac{u(x + r) - u(x - r)}{2}.$$

We recall that if we consider a sphere centered at x with radius r i.e. $S_r(x)$ then the probability of the arriving time is uniform. Therefore since u is represented by its expected value, we obtain the mean value theorem:

$$u(x) = \int_{S_r(x)} u \frac{1}{|S_r(x)|} = \frac{1}{|S_r(x)|} \int_{S_r(x)} u.$$

In the next section we briefly discuss how the tree methodology can be used for assets paying dividends.

6.8 Tree methods for dividend paying assets

In this section we discuss tree modifications to accommodate the case when the underlying asset pays continuous dividends, known discrete dividends as well as known cash dividends at a pre-specified time point t.

6.8.1 Options on assets paying a continuous dividend

Suppose an asset pays a dividend rate δ per unit of time. For example, suppose that the equity pays $\delta = 1\%$ annual dividend. If the original model used is geometric Brownian motion, applying Girsanov's theorem to (6.1), we obtain the stochastic equation:

$$dS_t = (r - \delta)S_t dt + \sigma S_t dW_t. \tag{6.35}$$

Note that equation (6.35) is exactly a geometric Brownian motion with the riskfree rate r replaced with $r - \delta$. Since everything stays the same, any tree constructed for a stochastic process with r for the drift will work by just replacing it with $r - \delta$. For example, the Trigeorgis tree in this case becomes:

$$\begin{aligned}
\Delta x &= \sqrt{\sigma^2 \Delta t + (r - \delta - \frac{1}{2}\sigma^2)^2 \Delta t^2} \\
p_u &= \frac{1}{2} + \frac{1}{2}\frac{(r - \delta - \frac{1}{2}\sigma^2)\Delta t}{\Delta x} \\
p_d &= 1 - p_u.
\end{aligned}$$

In our example $\delta = 0.01$. We can obviously write a routine for a general delta and in the case when there are no dividends just set $\delta = 0$. As an observation the unit time in finance is always 1 year. Therefore in any practical application if for instance expiry is one month we need to use $T = 1/12$.

6.8.2 Options on assets paying a known discrete proportional dividend

In this instance, the asset pays at some time τ in the future, a $\hat{\delta}$ dividend amount which is proportional to the stock value at that time τ. Specifically, the dividend amount is $\hat{\delta}S_\tau$. Due to the way the binomial tree is constructed we need a very simple modification to accommodate the dividend payment. We only care about the dividend payment if it happens during the option lifetime. If that is the case suppose $\tau \in [(i-1)\Delta t, i\Delta t]$.

When the dividend is paid the value of the asset drops by that particular amount. The justification is simple - the asset share value represents the value of the company. Since a ceratin amount is paid per share to the shareholders - an entity outside the company - the value of the company drops by that exact amount. To accommodate this we change the value of all nodes at time $i\Delta t$ by multiplying with $1-\hat{\delta}$ for a multiplicative tree in S_t and by adding $-\hat{\delta}$ for an additive tree. Specifically for multiplicative trees at node $(i,j) = (i\Delta t, S_j)$, the value is $S_0(1-\hat{\delta})u^j d^{i-j}$ where $u = e^{\Delta X_u}$ and $d = e^{\Delta X_d}$.

6.8.3 Options on assets paying a known discrete cash dividend

This is the most complex case and unfortunately this is the realistic case (most assets pay cash dividends). In this case, the major issue is that the tree becomes non-recombining after the dividend date. Suppose $\tau \in [(i-1)\Delta t, i\Delta t]$ and at that time the stock pays cash dividend D. The value of nodes after this time are supposed to be subtracted by D. Specifically at the node value (i,j) we have:

$$S_0(e^{\Delta X_u})^j(e^{\Delta X_d})^{i-j} - D.$$

At the next step the successor down will be $S_0(e^{\Delta X_u})^j(e^{\Delta X_d})^{i-j-1} - De^{\Delta X_d}$. On the other hand the successor up from the lower node: $S_0(e^{\Delta X_u})^{j-1}(e^{\Delta X_d})^{i-j-1} - D$ will be $S_0(e^{\Delta X_u})^j(e^{\Delta X_d})^{i-j-1} - De^{\Delta X_u}$. It is easy to see that the two nodes will not match unless $e^{\Delta X_u} = e^{\Delta X_d} = 1$ which makes the tree a straight line. Therefore, after step i the tree becomes non-recombining and for example at time $(i+m)\Delta t$ there will be $m(i+1)$ nodes rather than $i+m+1$ as in regular tree. That is a quadratic number of calculations that quickly become unmanageable.

The trick to deal with this problem is to make the tree non-recombining in the beginning rather than at the end. Specifically, we assume that S_t has 2 components i.e.

- $\tilde{S}_t$ a random component
- remainder depending on the future dividend stream:

$$\tilde{S}_t = \begin{cases} S_t & \text{when} \quad t > \tau \\ S_t - De^{-r(\tau-t)} & \text{when} \quad t \le \tau \end{cases}$$

Suppose $\tilde{S}_t$ follows a geometric Brownian motion with $\tilde{\sigma}$ constant. We calculate $p_u, p_d, \Delta X_u, \Delta X_d$ in the usual way with σ replaced with $\tilde{\sigma}$. The tree is constructed as before however the tree values are:

$$\text{at } (i,j), \text{ when} \quad t = i\Delta t < \tau : \tilde{S}_t(e^{\Delta X_u})^j(e^{\Delta X_d})^{i-j} + De^{-r(\tau-t)}$$
$$\text{at } (i,j), \text{ when} \quad t = i\Delta t \ge \tau : \tilde{S}_t(e^{\Delta X_u})^j(e^{\Delta X_d})^{i-j}.$$

This tree will be mathematically equivalent with a discretization of a continuous Brownian motion process that suddenly drops at the fixed time τ by the discrete amount D.

6.8.4 Tree for known (deterministic) time varying volatility

Suppose the stochastic process has a known time varying volatility $\sigma(t)$ and drift $r(t)$. That is the stock follows:

$$dS_t = r(t)S_t dt + \sigma(t)S_t dW_t$$

Fixing a time interval Δt, suppose at times $0, \Delta t, 2\Delta t, \ldots, \mu\Delta t$ the volatility values are $\sigma(i\Delta t) = \sigma_i, r(i\Delta t) = r_i$ then the corresponding drift term for the log process $X_t = \log S_t$ is $D_i = r_i - \frac{\sigma_i^2}{2}$. We keep the $\Delta X_u, \Delta X_d$ fixed (for example, we can take $\Delta X_u = \Delta X; \Delta X_d = -\Delta X$), this insures that the tree is recombining. However, we vary the probabilities at all steps so that the tree adapts to the time varying drift. We let the probabilities at time step i be denoted using $p_u^i = p_i$ and $p_d^i = 1 - p_i$. Thus, we need to have

$$\begin{cases} p_i\Delta x - (1-p_i)\Delta x & = D_i\Delta t_i \\ p_i\Delta x^2 + (1-p_i)\Delta x^2 & = \sigma_i^2\Delta t_i + D_i^2\Delta t_i^2 \end{cases} \tag{6.36}$$

Simplifying (6.36) we obtain

$$\begin{cases} 2p_i\Delta x - \Delta x & = D_i\Delta t_i \\ \Delta x^2 & = \sigma_i^2\Delta t_i + D_i^2\Delta t_i^2 \end{cases} \tag{6.37}$$

Equation (6.37) can be further simplified to:

$$p_i = \frac{1}{2} + \frac{D_i\Delta t_i}{2\Delta x} \tag{6.38}$$

Rearranging terms in (6.38) we obtain:

$$D_i^2\Delta t_i^2 + \sigma_i^2\Delta t_i - \Delta x^2 = 0 \tag{6.39}$$

Using the quadratic formula we solve for Δt_i to obtain

$$\Delta t_i = \frac{-\sigma_i^2 \pm \sqrt{\sigma_i^4 + 4D_i^2\Delta x^2}}{2D_i^2} \tag{6.40}$$

From (6.38), we have that

$$\Delta x^2 = \sigma_i^2\Delta t_i + D_i^2\Delta t_i^2 \tag{6.41}$$

Summing over i, we obtain

$$N\Delta x^2 = \sum_i^N \sigma_i^2\Delta t_i + \sum_i^N D_i^2\Delta t_i^2 \tag{6.42}$$

Dividing both sides by N we obtain:

$$\Delta x^2 = \frac{1}{N}\sum_{i}^{N}\sigma_i^2\Delta t_i + \frac{1}{N}\sum_{i}^{N}D_i^2\Delta t_i^2 \tag{6.43}$$

Taking $\overline{\Delta t} = \frac{1}{N}\sum_i^N \Delta t_i$ and solving for Δx we obtain:

$$\Delta x = \sqrt{\overline{\sigma^2}\,\overline{\Delta t} + \overline{D^2}\Delta t^2}$$

where $\overline{\sigma^2} = \frac{1}{N}\sum_{i=1}^N \sigma_i^2$ and $\overline{D^2} = \frac{1}{N}\sum_{i=1}^N D_i^2$. This finally gives us all the parameter values needed to construct the tree. The tree is approximately correct and the effects of the extra modification on estimation the value of the option is minimal.

6.9 Pricing path dependent options: Barrier Options

Barrier options are triggered by the action of the underlying asset hitting a prescribed value at some time before expiry. For example, as long as the asset remains below a predetermined barrier price during the whole life of the option, the contract will have a call payoff at expiry. Barrier options are clearly path dependent options. Path dependent options are defined as the right, but not the obligation, to buy or sell an underlying asset at a predetermined price during a specified time period, however the exercise time or price is dependent on the underlying asset value during all or part of the contract term. A path dependent option's payoff is determined by the path of the underlying asset's price.

There are two main types of barrier options:

1. The **In** type option, that exercises as specified in the contract as long as a particular asset level called barrier is reached before expiry. If the barrier is reached then the option is said to have knocked in. If the barrier level is not reached the option is worthless.

2. The **Out** type option, that becomes zero and is worthless if the Barrier level is not reached. If the barrier is reached then the option is said to have knocked out.

We further characterize the barrier option by the position of the barrier relative to the initial value of the underlying:

1. If the barrier is above the initial asset value, we have an Up option.

2. If the barrier is below the initial asset value, we have a Down option.

We note that if an Out option starts as regular American and barrier is hit option becomes zero and similarly that an In option typically starts worthless and if the barrier is hit the option becomes a regular American option.

Given a barrier B and path $= S_{t1}, \ldots, S_{tN}$, the terminal payoff of the barrier options can be written as:

1. Down and Out Call $(S_T - K)_+ 1_{\{\min(S_{t1},\ldots,S_{tN})>B\}}$
2. Up and Out Call $(S_T - K)_+ 1_{\{\max(S_{t1},\ldots,S_{tN})<B\}}$
3. Down and In Call $(S_T - K)_+ 1_{\{\min(S_{t1},\ldots,S_{tN})\leq B\}}$
4. Up and In Call $(S_T - K)_+ 1_{\{\max(S_{t1},\ldots,S_{tN})\geq B\}}$
5. Down and Out Put $(K - S_T)_+ 1_{\{\min(S_{t1},\ldots,S_{tN})>B\}}$
6. Up and Out Put $(K - S_T)_+ 1_{\{\max(S_{t1},\ldots,S_{tN})<B\}}$
7. Down and In Put $(K - S_T)_+ 1_{\{\min(S_{t1},\ldots,S_{tN})\leq B\}}$
8. Up and In Put $(K - S_T)_+ 1_{\{\max(S_{t1},\ldots,S_{tN})\geq B\}}$

Remark 6.9.1. *Generally, the In type options are much harder to solve than Out type options. However, it is easy to see that we have the following Put-Call parity relations.*

Down and Out Call (K, B, T) + *Down and In Call* $(K, B, T) =$ *Call* (K, T) *for all* B.

This is easy to prove by looking at the payoffs and remarking that:

$$1_{\min(S_{t1},\ldots,S_{tN})>B} + 1_{\min(S_{t1},\ldots,S_{tN})\leq B} = 1,$$

and recalling that option price at t is the discounted expectation of the final payoff. Therefore applying expectations will give exactly the relation needed. Thus one just needs to construct a tree method for Down type options and use the In-Out Parity above to obtain the price of an In type option.

6.10 Trinomial tree method and other considerations

The trinomial trees provide an effective method of numerical calculation of option prices within the Black-Scholes model. Trinomial trees can be built in a similar way to the binomial tree. To create the jump sizes u and d and the transition probabilities p_u and p_d in a binomial tree model we aim to match these parameters to the first two moments of the distribution of our geometric Brownian motion. The same can be done for our trinomial tree for u, d, p_u, p_m, p_d.

For the trinomial tree, one cannot go up and down by different amounts to keep it recombining. The condition imposed is the fact that the discrete increment needs to match the continuous once that is:

$$E(\Delta x) = \Delta x p_u + o p_u + (-\Delta x) p_d = D dt, \quad D = r - \frac{\sigma^2}{2}$$

$$E(\Delta x^2) = \Delta x^2 p_u + o p_u + (-\Delta x)^2 p_d = \sigma^2 \Delta t + D^2 \Delta t^2$$

$$p_u + p_m + p_d = 1$$

These are 3 equations and 4 unknowns. So it has an infinite number of solutions. However we have that:

$$\begin{aligned} p_u &= \frac{1}{2}\left(\frac{\sigma^2 \Delta t + D^2 \Delta t^2}{\Delta x^2} + \frac{D \Delta t}{\Delta x}\right) \\ p_m &= 1 - \frac{\sigma^2 \Delta t + D^2 \Delta t^2}{\Delta x^2} \\ p_d &= \frac{1}{2}\left(\frac{\sigma^2 \Delta t + D^2 \Delta t^2}{\Delta x^2} - \frac{D \Delta t}{\Delta x}\right) \end{aligned}$$

These probabilities need to be numbers between 0 and 1. Imposing this condition we obtain a sufficient condition: $\Delta x \geq \sigma \sqrt{3 \Delta t}$. This condition will be explained later in this section. Any Δx with this property produces a convergent tree.

The trinomial tree is an alternate way to approximate the stock price model. The stock price once again follows the equation:

$$dS_t = r S_t dt + \sigma S_t dW_t. \tag{6.44}$$

In [38] the authors work with a continuously paying dividend asset and the drift in the equation is replaced by $r - \delta$. All the methods we will implement require an input r. One can easily obtain the formulas for a continuously paying dividend asset by just replacing this parameter r with $r - \delta$. Once again, it is equivalent to work with the return $X_t = \log S_t$ instead of directly with the stock and we obtain:

$$dX_t = \nu dt + \sigma dW_t, \qquad \text{where } \nu = r - \frac{1}{2}\sigma^2. \tag{6.45}$$

The construction of the trinomial tree is equivalent to the construction of the binomial tree described in previous sections. A one step trinomial tree is presented below:

$$\begin{array}{cc} & S_u \\ S & S \\ & S_d \end{array}$$

Trinomial trees allows the option value to increase, decrease or remain stationary at every time step as illustrated above.

Once again we match the expectation and variance. In this case the system contains three equations and three unknowns so we do not have a free choice as in the binomial tree case. In order to have a convergent tree, numerical experiments have shown that we impose a condition such that

$$\Delta x \geq \sigma\sqrt{3\Delta t} \tag{6.46}$$

Hence for stability there must be restrictions on the relative sizes of Δx and Δt. The condition ensures that our method is stable and converges to the exact solution. Please refer to [38] for details about the stability condition. Once the tree is constructed we find an American or European option value by stepping back through the tree in a similar manner with what we did for the binomial tree. The only difference is that we calculate the discounted expectation of three node values instead of two as we did for the binomial tree. The main advantage of the trinomial tree over the binomial tree construction is the fact that the trinomial tree is appropriate for pricing options when the volatility varies. This is because when we vary the volatility, we introduce other parameters in the model that increases the number of equations which is easily solved by the trinomial tree construction method.

The trinomial tree produces more paths (3^n) than the binomial tree (2^n). Surprisingly the order of convergence is not affected for this extra number. In both cases the convergence of the option values is of the order $O(\Delta x^2 + \Delta t)$. Optimal convergence is always guaranteed due to condition(6.46).

Condition (6.46) makes a lot of difference when we deal with barrier options i.e. options that are path dependent. This is due to the fact that the trinomial tree contains a larger number of possible nodes at each time in the tree. The trinomial tree is capable of dealing with the situations when the volatility changes over time i.e., is a function of time. For a detailed example of recombining trinomial tree for valuing real options with changing volatility, please refer to [**?**]. In that study, the trinomial tree is constructed by choosing a parameterization that sets a judicious state space while having sensible transition probabilities between the nodes. The volatility changes are modeled with the changing transition probabilities while the state space of the trinomial tree is regular and has a fixed number of time and underlying asset price levels.

What is the meaning of little o and big O?

Suppose we have two functions f and g. We say that f is of order little o of g at x_0:

$$f \sim o(g) \Leftrightarrow \lim_{x \to x_0} \frac{f(x)}{g(x)} = 0$$

We say that f is of order big O of g at x_0:

$$f \sim O(g) \Leftrightarrow \lim_{x \to x_0} \frac{f(x)}{g(x)} = C$$

where C is a constant.

In our context if we calculate the price of an option using an approximation (e.g., trinomial tree) call it $\hat{\Pi}$, and the real (unknown) price of the option call it Π, then we say that the approximation is of the order $O(\Delta x^2 + \Delta t)$ and we mean that:

$$|\hat{\Pi} - \Pi| = C(\Delta x^2 + \Delta t)$$

whenever Δx and Δt both go to zero for some constant C.

6.11 Problems

1. (The Martingale problem on L). Let (f, g) be a pair of functions in L. Find a process $\{X_t\}_t$ defined on E such that:

$$M_t = f(X_t) - f(X_0) - \int_0^t g(X_0)ds$$

is a martingale.

2. Construct a binomial tree ($u = \frac{1}{d}$) with three quarters for an asset with present value \$100. If $r = 0.1$ and $\sigma = 0.4$, using the tree compute the prices of:

   (a) An European Call option with strike 110.

   (b) An European Put option with strike 110.

   (c) An American Put option with strike 110.

   (d) An European Call option with strike 110, assuming that the underlying asset pays out a dividend equivalent to 1/10 of its value at the end of the second quarter.

   (e) An American Call option with strike 110 for an underlying as in iv).

   (f) An *asian option*, with strike equal to the average of the underlying asset values during the period.

   (g) A *barrier option* (for different barriers).

3. In the binomial model obtain the values of u, d and p given the volatility σ and the risk free interest rate r, for the following cases:

   (a) $p = \frac{1}{2}$.

   (b) $u = \frac{1}{d}$. **Hint**: $E\left(S_{t+\Delta t}^2\right) = S_t^2 e^{(2r+\sigma^2)\Delta t}$

4. Compute the value of an option with strike \$100 expiring in four months on underlying asset with present value by \$97, using the binomial model. The risk free interest rate is 7% per year and the volatility is 20%. Assume that $u = \frac{1}{d}$.

5. In a binomial tree with n steps, let $f_j = f_{u\ldots ud\ldots d}$ (j times u and $n - j$ times d). For an European Call option expiring at $T = n\Delta t$ with strike K, show that

$$f_j = \max\{Su^j d^{n-j}, 0\}$$

In particular, if $u = \frac{1}{d}$, we have that

$$f_j \geq 0 \Longleftrightarrow j \geq \frac{1}{2}(n + \frac{log(K/S)}{log(u)})$$

(a) Is it true that the probability of positive payoff is the probability of $S_T \geq K$? Justify your answer.

(b) Find a general formula for the present value of the option.

6. We know that the present value of a share is \$40, and that after one month it will be \$42 or \$38. The risk free interest rate is 8% per year continuously compounded.

   (a) What is the value of an European Call option that expires in one month, with strike price \$39?

   (b) What is the value of a Put option with the same strike price?

7. Explain the difference between pricing an European option by using a binomial tree with one period and assuming no arbitrage, and by using risk neutral valuation.

8. The price of a share is \$100. During the following six months the price can go up or down in a 10% per month. If the risk free interest rate is 8% per year, continuously compounded.

   (a) what is the value of an European Call option expiring in one year with strike price \$100?

   (b) Compare with the result obtained when the risk free interest rate is monthly compounded.

9. The price of a share is \$40, and it is incremented in 6% or it goes down in 5% every three months. If the risk free interest rate is 8% per year, continuously compounded, compute:

   (a) The price an European Put option expiring in six months with a strike price of \$42.

   (b) The price of an American Put option expiring in six months with a strike price of \$42.

10. The price of a share is \$25, and after two months it will be \$23 or \$27. The risk free interest rate is 10% per year, continuously compounded. If S_T is the price of the share after two months, what is the value of an option with the same expiration date (i.e. two months) and payoff S_T^2?

11. Calculate the price for a forward contract, that is: there exists the obligation to buy at the expiration. Hint: construct a portfolio so that taking it back on time, it is risk free.

12. The price of a share is \$ 40. If $\mu = 0.1$ and $\sigma^2 = 0.16$ per year, find a 95 % confidence interval for the price of the share after six months (i.e. an interval $I_{0.95} = (\underline{S}, \overline{S})$ so that $p(S \in I_{0.95}) = 0.95$). Hint: use that if Z is a stochastic variable with standard normal distribution, then $p(-\,1.96 \leq Z \leq 1.96) \simeq 0.95$).

13. Suppose that the price of a share verifies that $\mu = 16\%$ and the volatility is 30%. If the closing price of the share at a given day is \$ 50, compute:

    (a) The closing expected value of the share for the next day.

    (b) The standard deviation of the closing price of the share for the next day.

    (c) A 95 % confidence interval for the closing price of the share for the next day.

14. American Down and Out Call.
    Given $K = 100, T = \frac{1}{30}, S = 100, \sigma = 0.6, r = 0.01$ and $B = 95$. Construct a tree for all nodes below the barrier value $= 0$.

15. Consider the parabolic equation,

$$\begin{cases} \frac{\partial u}{\partial t} + \alpha\left(x,t\right)\frac{\partial^2 u}{\partial x^2} + \beta\left(x,t\right)\frac{\partial u}{\partial x} = 0; \alpha > 0 \\ u\left(x,T\right) = \phi\left(x\right) \end{cases}$$

    Using a convenient change of variable transform equation (15) into the 1-dimensional heat equation.