Chapter I: Introduction to Stochastic Process

Liu Yanbo

May 24, 2018

Abstract

The aim of this chapter is to get you guys be familiar with quantitative tools in discrete-time stochastic process and their applications in dynamic programming methods. This chapter is based on the following materials.

- Durrett, Rick. Probability: theory and examples. Cambridge university press, 2010.
- Stokey, Nancy L. Recursive methods in economic dynamics. Harvard University Press, 1989.
- Ross, Sheldon M. Stochastic processes. Vol. 2. New York: John Wiley & Sons, 1996.
- Lecture Notes of "Dynamic Optimization" by Professor Zhu Shenghao, https://shenghaozhu.weebly.com/teaching-materials.html

1 Introduction to Stochastic Process

A stochastic process is a collection of random variables indexed by time.

An alternate view is that it is a probability distribution over a space of paths; this path often describes the evolution of some random value, or system, over time. In a deterministic process, there is a fixed trajectory (path) that the process follows, but in a stochastic process, we do not know a priori which path we will be given. One should not regard this as having no information of the path since the information on the path is given by the probability distribution. For example, if the probability distribution is given as one path having probability one, then this is equivalent to having a deterministic process. Also, it is often interpreted that the process evolves over time. However, from the formal mathematical point of view, a better picture to have in mind is that we have some underlying (unknown) path, and are observing only the initial segment of this path.

For example, the function $f : \mathbb{R}_{\geq} \to \mathbb{R}$ given by $f(t) = t$ is a deterministic process, but a "random process" $f : \mathbb{R}_{\geq} \to \mathbb{R}$ given by $f(t) = t$ with probability $\frac{1}{2}$ and $f(t) = -t$ with probability $\frac{1}{2}$ is a stochastic process. This is a rather degenerate example and we will later see more examples of stochastic processes.

We are still dealing with a single basic experiment that involves outcomes governed by a probability law. However, the newly introduced time variable allows us to ask many new interesting questions. We emphasize on the following topics:

- We tend to focus on the dependencies in the sequence of values generated by the process. For example, how do future prices of a stock depend on past values?
- We are interested in long-term averages involving the entire sequence of generated values. For example, what is the fraction of time that a machine is idle?
- We are interested in boundary events. For example, what is the probability that within a given hour all circuits of some telephone system become simultaneously busy

A stochastic process has discrete-time if the time variable takes positive integer values, and continuous-time if the time variable takes positive real values. We start by studying discrete time stochastic processes. These processes can be expressed explicitly, and thus are more 'tangible', or 'easy to visualize'. Later we address continuous time processes.

2 Simple Random Walk

Let $Y_1, Y_2, ...$ be i.i.d random variables such that $Y_i = 1$ or $Y_i = -1$. Let $X_0 = 0$ and

$$X_k = Y_1 + ... + Y_k$$

for all $k \geq 1$. This gives a probability distribution over the sequences $\{X_0, X_1, ..., \}$, and thus defines a discrete time stochastic process. This process is known as the one-dimensional simple random walk, which we conveniently refer to as random walk from now on.

By the central limit theorem, we know that for large enough n, the distribution of $\frac{1}{\sqrt{n}}X_n$ converges to the normal distribution with mean 0 and variance 1. This already tells us some information about the random walk. We state some further properties of the random walk.

Theorem 2.1.

- *$\mathbb{E}[X_k] = 0$ for all k;*
- *(Independent increment) For all $0 = k_0 \leq k_1 \leq ... \leq k_r$, the random variable $X_{k_{i+1}} - X_{k_i}$, for $0 \leq i \leq r-1$ are mutually independent.*
- *(Stationary) For all $h \geq 1$ and $k \geq 0$, the distribution of $X_{k+h} - X_k$ is the same as the distribution of X_h*

Example 2.2.

- *Suppose that a gambler plays the following game. At each turn the dealer throws an un-biased coin, and if the outcome is head the gambler wins \$1, while if it is head she loses \$1. If each coin toss is independent, then the balance of the gambler has the distribution of the simple random walk.*

- *Random walk can also be used as a (rather inaccurate) model of stock price.*

3 Markov Chain

One important property of the simple random walk is that the effect of the past on the future is summarized only by the current state, rather than the whole history. In other words, the distribution of X_{k+1}depended only on the value of X_k, not on the whole set of values of $X_0, X_1, X_2, ..., X_k$. A stochastic process with such property is called a *Markov chain.*

More formally, let $X_0, X_1, ...,$ be a discrete time stochastic process where each X_i takes value in some discrete set S (note that this is not the case in the simple random walk). The set S is called the *state space.* We say that the random process has the *Markov property* if

$$\mathbb{P}(X_{n+1} = i | X_n, X_{n-1}, ..., X_0) = \mathbb{P}(X_{n+1} = i | X_n)$$

for all $n \geq 0$ and $i \in S$. We will discuss the case when S is a finite set. In this case, we let $S = [m]$ for some positive integer m.

A stochastic process with the Markov property is called a Markov chain. Note that a finite Markov chain can be described in terms of the transition probabilities

$$p_{ij} = \mathbb{P}(X_{n+1} = j | X_n = i), i, j \in S$$

One could easily see that

$$\sum_{j \in S} p_{ij} = 1, \forall i \in S$$

Thus all the elements of a Markov chain could be encoded into a transition probability matrix

$$A = \begin{Bmatrix} p_{11} & p_{21} & \cdots & p_{m1} \\ p_{12} & p_{22} & \cdots & p_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ p_{1m} & p_{2m} & \cdots & p_{mm} \end{Bmatrix}$$

Note that the sum of each column is equal to one.

Example 3.1. • *A machine can be either working or broken on a given day. If it is working, it will break down in the next day with probability 0.01, and will continue working with probability 0.99. If it breaks down on a given day, it will be repaired and be working in the next day with probability 0.8, and will continue to be broken down with probability 0.2. We can model this machine by a Markov chain with two states: working, and broken down. The transition probability matrix is given by*

$$\begin{bmatrix} 0.99 & 0.8 \\ 0.01 & 0.2 \end{bmatrix}$$

- *A simple random walk is an example of a Markov chain. However, there is no transition probability matrix associated with the simple random walk since the sample space is of infinite cardinatlity.*
  Let $r_{ij}(n) = \mathbb{P}(X_n = j|X_0 = i)$ *be the n-th transition probabilities. These probabilities satisfy the recurrence relation*

  $$r_{ij}(n) = \sum_{k=1}^{m} r_{ik}(n-1)p_{kj}$$

  for $n > 1$*, and* $r_{ij}(1) = p_{ij}$*. Hence the n-step transition probability matrix could be easily shown to be* A^n*.*
  A stationary distribution of a Markov chain is a probability distribution over the state space $(where \mathbb{P}(X_0 = j) = \pi_j)$ *such tat*

  $$\pi_j = \sum_{k=1}^{m} \pi_k \cdot p_{kj}$$

  for $\forall j \in S$

- *Let* $S = \mathbb{Z}_n$ *and* $X_0 = 0$*. Consider the Markov chain* $X_0, X_1, ...$ *such that* $X_{n+1} = X_n + 1$ *with probability* $\frac{1}{2}$ *and* $X_{n+1} = X_n - 1$ *with probability* $\frac{1}{2}$*. Then the stationary distribution of this Markov chain is* $\pi_i = \frac{1}{n}$ *for* $\forall i$*.*

Here is one fundamental theorem for Markov chain process, which is as,

Theorem 3.2. *If* $p_{ij} > 0$ *for all* $i, j \in S$ *then there exists a unique stationary distribution of the system. Moreover,*

$$lim_{n \to \infty} r_{ij}(n) = \pi_{ij}$$

for $\forall i, j \in S$*. However, a corresponding theorem is not true if we consider infinite state spaces.*

4 Martingale

Definition 4.1. *A discrete-time stochastic process* $\{X_0, X_1, ...\}$ *is a martingale if*

$$X_t = \mathbb{E}\left[X_{t+1}|\mathcal{F}_t\right]$$

for all $t \geq 0$*, where* $\mathcal{F}_t = \{X_0, ...X_t\}$ *(hence we are conditioning on the initial segment of the process).*

This says that our expectated gain in the process is zero at all times. We can also view this definition as a Mathematical formalization of a game of chance being fair.

Theorem 4.2. *For all* $t \geq s$*, we have* $X_s = \mathbb{E}\left[X_t|\mathcal{F}_s\right]$

Proof: This easily follows from deduction and iterated law of expectations

Example 4.3. • *Random walk is a martingale*

- *The balance of a roulette player is not a martingale (we always have* $X_k > \mathbb{E}[X_{k+1}|\mathcal{F}_t]$*)*
- *Let* $Y_1, Y_2, ...$ *be iid random variable such that* $Y_i = 2$ *with probability* $\frac{1}{3}$ *and* $Y_i = \frac{1}{2}$ *with probability* $\frac{2}{3}$*. Let* $X_0 = 0$*,* $X_k = \prod_{i=1}^{k} Y_i$*. Then* $\{X_0, X_1, ...\}$ *forms a martingale.*

5 Poisson Process

To state the deÖnition of a Poisson process we use the deÖnition of a counting process. I copy the following two concepts from Ross (1983) textbook, *Stochastic Processes.*

Definition 5.1. *A stochastic process* $\{N(t), t \geq 0\}$ *is said to be a counting process if* $N(t)$ *represents the total number of "event" that have occurred up to time t. Hence, a counting process* $N(t)$ *must satisfy:*

- $N(t) \geq 0$
- $N(t)$ *is integer valued*
- *If* $s < t$*, then* $N(s) \leq N(t)$
- *For* $s < t$*,* $N(t) - N(s)$ *equals the number of events that have occurred in the interval* $(s, t]$

Definition 5.2. *The counting process* $\{N(t), t \geq 0\}$ *is said to be a Poisson process having rate* λ*,* $\lambda > 0$*, if:*

- $N(0) = 0$
- *The process has independent increments*
- *The number of events in any interval of length t is Poisson distributed with mean* λt*. That is, for all s, t≥0,*

$$\mathbb{P}\{N(t+s) - N(t) = n\} = e^{-\lambda t}\frac{(\lambda t)^n}{n!}$$

  for $\forall n = 0, 1, 2, 3, ...$

The Poisson process is a collection $\{N(t), t \geq 0\}$ of random variables, where $N(t)$ is the number of events that have occurred up to time t (starting from time 0). The number of events between time a and time b is given as $N(b) - N(a)$ and has a Poisson distribution. Each realization of the process $\{N(t)\}$ is a non-negative integer-valued step function that is non-decreasing, but for intuitive purposes it is usually easier to think of it as a point pattern on $[0, \infty)$ (the points in time where the step function jumps, i.e. the points in time where an event occurs).

6 Dynamic Optimization with a Jump Process

Here we plan to solve an optimization problem,

$$max \mathbb{E}_0 \int_0^T f(t,x,u)dt + \theta(x(T),T)$$

$$s.t. dx = g(t,x,u)dt + \sigma(t,x,u)dz + m(t,x,u,\omega)dq$$

$$x(0) = \overline{x}$$

where $z(t)$ is a standard Brownian motion and $q(t)$ is a Poisson process. The term of $m(t,x,u,\omega)dq$ means that when a jump happens, x changes from $x(t)$ to $x(t) + m(t,x,u,\omega)$, i.e. the jump part dominates the diffusion part. If no jump happens, x follows $dx = g(t,x,u)dt + \sigma(t,x,u)dz$. Here the jump size $m(t,x,u,\omega)$ is a random variable. I use ω to represent sample path. We usually use a Poisson process to represent a jump process.

6.1 Principle of Dynamic Optimization

To solve the dynamic optimization problem, we could use Ito formula for stochastic process with jumps. But now we will the recursive structure of the stochastic process with jumps directly to derive Hamilton-Jacobian-Bellman equation. The optimal value function is as,

$$max \mathbb{E}_0 \int_0^T f(t,x,u)dt + \theta(x(T),T)$$

$$s.t. dx = g(t,x,u)dt + \sigma(t,x,u)dz + m(t,x,u,\omega)dq$$

Let us derive HJB as

$$\begin{aligned}
J(t,x) &= max_u \mathbb{E}_t \int_t^T f(s,x,u)ds + \theta(x(T),T) \\
&= max_u \mathbb{E}_t \int_t^{t+\Delta t} f(s,x,u)ds + \int_{t+\Delta t}^T f(s,x,u)ds + \theta(x(T),T) \\
&= max_u \mathbb{E}_t (\int_t^{t+\Delta t} f(s,x,u)ds + max_u \mathbb{E}_{t+\Delta t} \int_{t+\Delta t}^T f(s,x,u)ds + \theta(x(T),T)) \\
&= max_u \mathbb{E}_t (\int_t^{t+\Delta t} f(s,x,u)ds + J(t+\Delta t, x+\Delta x)) \\
&= max_u \mathbb{E}_t (f(\widetilde{t}, x(\widetilde{t}), u(\widetilde{t}))\Delta t + J(t+\Delta t, x+\Delta x)) \\
&= max_u \mathbb{E}_t (f(\widetilde{t}, x(\widetilde{t}), u(\widetilde{t}))\Delta t + \lambda \Delta t J(t+\Delta t, x+m) \\
&+ (1-\lambda\Delta t)(J(t,x) + J_t \Delta t + J_x \Delta x + \frac{1}{2} J_{xx}(\Delta x)^2)) \\
&= max_u \mathbb{E}_t (f(\widetilde{t}, x(\widetilde{t}), u(\widetilde{t}))\Delta t + \lambda J(t+\Delta t, x+m)\Delta t + (1-\lambda\Delta t)(J(t,x) \\
&+ J_t \Delta t + J_x g \Delta t + J_x \sigma \Delta z + \frac{1}{2} J_{xx} \sigma^2 \Delta t))
\end{aligned} \tag{1}$$

Note that $\mathbb{E}_t \Delta t = 0$. Taking expectation operator,we have

$$J(t,x) = max_u(f(\widetilde{t}, x(\widetilde{t}), u(\widetilde{t}))\Delta t + \lambda \mathbb{E}_t J(t+\Delta t, x+m)\Delta t + (1-\lambda\Delta t)(J(t,x) + J_t\Delta t + J_x g\Delta t + \frac{1}{2}J_{xx}\sigma^2\Delta t))$$

We have the term of $\mathbb{E}_t J(t + \Delta t, x + m)$ in the equation, because $m(t, x, u, m)$ is a random variable.
Dividing by Δt on both sides and letting $\Delta t \to 0$, we have

$$\lambda J - J_t = max_u(f(t,x,u) - \lambda \mathbb{E}_t J(t, x+m) + J_x g + \frac{1}{2}J_{xx}\sigma^2)$$

Example 6.1. (Investment return with a jump process)
Now the agent can access a risky asset: a stock with possibility to receive dividend. The value of the stock follows

$$dS(t) = \alpha S(t)dt + \sigma S(t)dz(t) + \theta S(t)dq(t)$$

The jump size θ could be a random variable. Here for simplicity, I assume θ is a constant. When a jump happens, the value of stock changes from $S(t)$ to $(1+\theta)S(t)$. The agent chooses optimal consumption and investment rules, $c(t)$ and $\omega(t)$ to maximize the utility

$$max_{c(t),\omega(t)} \mathbb{E}_0 \int_0^\infty e^{-\beta t} \frac{(c(t))^{1-\gamma}}{1-\gamma} dt$$

$s.t., dw(t) = (r\omega(t) + (\alpha - r)\omega(t)w(t) - c(t))dt + \sigma\omega(t)w(t)dz(t) + \theta\omega(t)w(t)dq(t)$
Let,

$$V(w(t)) = max_{c(s),\omega(s)} \mathbb{E}_t \int_t^\infty e^{-\beta(s-t)} \frac{(c(s))^{1-\gamma}}{1-\gamma} ds$$

Then HJB is $(\lambda + \beta)V(w) = max_{c(t),\omega(t)}(\frac{(c(t))^{1-\gamma}}{1-\gamma} + \lambda V((1 + \theta\omega(t))w(t)) + V^{'}(w)(rw(t) + (\alpha - r)\omega(t)w(t) - c(t)) + \frac{1}{2}V^{''}(w)\sigma^2\omega^2(t)w^2(t))$
Then first order conditions are

$$(c(t))^{-\gamma} = V^{'}(w)$$

$$\lambda V^{'}((1 + \theta\omega(t))w(t))\theta w(t) + V^{'}(w)(\alpha - r)w(t) + V^{''}(w)\sigma^2\omega(t)w^2(t) = 0$$

Guess

$$V(w(t)) = \frac{A}{1-\gamma}(w(t))^{1-\gamma}$$

Thus

$$c(t) = A^{-\frac{1}{\gamma}} w(t)$$
$$\omega(t) = \overline{\omega}$$

Where $\overline{\omega}$ solves

$$\gamma\sigma^2\overline{\omega} - \lambda\theta(1 + \theta\overline{\omega}^{-\gamma}) - (\alpha - r) = 0$$

From HJB we find that

$$A = (\frac{\lambda + \beta - \lambda(1 + \theta\overline{\omega})^{1-\gamma} - (1-\gamma)(r + (\alpha - r)\overline{\omega} - \frac{1}{2}\gamma\sigma^2\overline{\omega}^2)}{\gamma})^{-\gamma}$$

Chapter II: Stochastic Calculus

Liu Yanbo

May 24, 2018

Abstract

This part of is intended to prepare you guys for the continuous-time models both in *Macroeconomics II* and *Econometrics II*. Among all five chapters of Math IIB, stochastic calculus is definitely the hardest part. This chapter is based on the following materials.

- Øksendal, Bernt. 'Stochastic differential equations.' Stochastic differential equations. Springer Berlin Heidelberg, 2003. 65-84.
- Karatzas, Ioannis, and Steven Shreve. Brownian motion and stochastic calculus. Vol. 113. Springer Science & Business Media, 2012.
- Shreve, Steven E. Stochastic calculus for finance II: Continuous-time models. Vol. 11. Springer Science & Business Media, 2004.
- Lecture Notes of "Introduction to Financial Mathematics II", National Chiao Tung University, http://ocw.nctu.edu.tw/
- Lecture Notes of "Dynamic Optimization" by Professor Zhu Shenghao, https://shenghaozhu.weebly.com/teaching-materials.html

1 Continuous-Time Martingales

1.1 Stochastic Processes

Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a probability space and let $I \subseteq [0, \infty)$ be an interval.

Definition 1.1. *A real-valued stochastic process $X = (X_t)_{t \in I}$ is a family of random variables $(X_t : t \in I)$ on $(\Omega, \mathcal{F})$*

Remark 1.2. *We may regard the stochastic process X as a function of two random variables*

$$X : I \times \Omega \to \mathcal{R}$$

$$(t, \omega) \to X_t(\omega))$$

so we could treat the random process as the two-dimensional function

- *For fixed $\omega \in \Omega$, then*
  $$t \to X_t(\omega)$$
  is a function: $I \to \mathcal{R}$, which is called a path of X
- *For fixed $t \in I$, then*
  $$\omega \to X_t(\omega)$$
  is a function: $\Omega \to \mathcal{R}$, which is a random variable.

1.2 Martingales in continuous time

Here Let $I = [0, \infty)$

Definition 1.3. *A stochastic process $X = (X_t)_{t \leq 0}$ is called a martingale with respect to $\mathbb{P}$if,*

- *for $0 \leq s \leq t < \infty$,*
$$\mathbb{E}\left[X_t | \mathcal{F}_s\right] = X_s$$
  with $\mathbb{P} - a.s.$

Definition 1.4. *A stochastic process $X = (X_t)_{t \leq 0}$ is called a submartingale with respect to $\mathbb{P}$ if,*

- *for $0 \leq s \leq t < \infty$,*
$$\mathbb{E}\left[X_t | \mathcal{F}_s\right] \geq X_s$$
  with $\mathbb{P} - a.s.$

Definition 1.5. *A stochastic process $X = (X_t)_{t \leq 0}$ is called a supermartingale with respect to $\mathbb{P}$ if,*

- *for $0 \leq s \leq t < \infty$,*
$$\mathbb{E}\left[X_t | \mathcal{F}_s\right] \leq X_s$$
  with $\mathbb{P} - a.s.$

Example 1.6. *Let $Z \in L^1(\mathbb{P})$, then the process (X_t) defined by*

$$X_t = \mathbb{E}\left[Z | \mathcal{F}_t\right]$$

is one martingale, and this is easily checked by using the iterated law of expectations and Fubini theorem

1.3 Doob-Meyer Decomposition

Theorem 1.7. (Doob-Meyer Decomposition) Let $X = (X_t)_{t \geq 0}$ be a supermartingale, then X admits a unique decomposition

$$X_t = X_0 + M_t - A_t$$

where M is a martingale with $M_0 = 0$ and A is an increasing, right-continuous previsible process with $A_0 = 0$

Corollary 1.8. *Let $M \in \mathcal{M}^2$ be right-continuous. Then there exists a unique right-continuous previsible process $\langle M \rangle = (\langle M \rangle_t)_{t \geq 0}$ with $\langle M \rangle_0 = 0$ such that the process $M^2 - \langle M \rangle$ is a martingale.*

Remark 1.9. *M is martingale $\Rightarrow M^2$ is submartingale $\Rightarrow -M^2$ is supermartingale, since $\mathbb{M}\left[M_t^2 | \mathcal{F}_s\right] \geq \mathbb{M}\left[M_t | \mathcal{F}_s\right]^2 = M_s^2 \Rightarrow$ Apply the Doob-Meyer decompostion to $-M^2$, and the A_t in the theorem is what we call $\langle M \rangle$*

Definition 1.10. *$\langle M \rangle_0 = 0$ is called the quadratic variation of M*

Theorem 1.11. *Let $M \in \mathcal{M}^{2,c}$. For partition Π of $[0,t]$, set*

$$\|\Pi\| \doteq max_{1\leq k\leq m}|t_k - t_{k-1}|$$

we have,

$$plim_{\|\Pi\|\to 0}\sum_{k=1}^{m}|M_{t_k} - M_{t_{k-1}}|^2 = \langle M\rangle_t$$

i.e., for any $\epsilon > 0, \theta > 0$, there exists $\delta > 0$ such that,

$$max_{1\leq k\leq m}|t_k - t_{k-1}| < \delta \Rightarrow \mathbb{P}(||M_{t_k} - M_{t_{k-1}}|^2 - \langle M\rangle_t| > \epsilon) < \theta$$

2 Brownian Motions

2.1 Scaled Random Walk

Before introducing Brownian Motions, let us discuss symmetric random walk at first,

Exercise 2.1. (construction of a symmetric random walk) Toss the coin, and the probabilities for head and tail are equal, which is as,

$$Prob(Head) = Prob(Tail) = \frac{1}{2}$$

The successive outcome of the toss $\omega = \omega_1\omega_2\omega_3\omega_4...\omega_n...$, where ω_n is the outcome of the nth toss. The sample space Ω is given by

$$\Omega = \{\omega : \omega = \omega_1\omega_2..., \omega_i =' H' or' T'\}$$

Let

$$X = \begin{cases} 1, & if \quad \omega_n = H \\ -1, & if \quad \omega_n = T \end{cases}$$

and $(X_n)_{n\geq 1}$ is independent

Definition 2.2. (symmetric random walk) Define

$$M_0 = 0$$

$$M_k = \sum_{i=1}^{k} X_i, k = 1, 2, 3, 4...$$

The process $(M_k)_{k\geq 0}$ is a symmetric random walk

Theorem 2.3. *A random walk has independent increments, i.e., any $0 = t_0 < t_1 < t_2 < ... < t_m = t$,$(t_i \in \mathbb{N})$, the increments of the random walk*

$$M_{t_1}, M_{t_2} - M_{t_1}, M_{t_3} - M_{t_2}, ..., M_{t_m} - M_{t_{m-1}}, ...$$

are independent.

Proof follows the definition, and is omitted here.

Theorem 2.4. *The random variable*

$$M_{t_k} - M_{t_{k-1}} = \sum_{i=t_{k-1}+1}^{t_k} X_i$$

has expectation 0 and variance $t_k - t_{k-1}$

Proof: Since

$$\mathbb{E}[X_i] = 0$$
$$Var(X_i) = \mathbb{E}\left[X_i^2\right] - (\mathbb{E}[X_i])^2 = 1$$

we have

$$\mathbb{E}\left[M_{t_k} - M_{t_{k-1}}\right] = \sum_{i=t_{k-1}+1}^{t_k} \mathbb{E}[X_i] = 0$$

$$Var(M_{t_k} - M_{t_{k-1}}) = \sum_{t_{k-1}+1}^{t_k} Var(X_i) = t_k - t_{k-1}$$

and the independence is due to theorem 2.3

Theorem 2.5. (M_k) *is a martingale with respect to* $(\mathcal{F}_k^X)$

Proof: Since $(X_n)_{n\geq 1}$ is independent,

$$\mathbb{E}\left[M_k - M_{k-1}|\mathcal{F}_{k-1}^X\right] = \mathbb{E}\left[X_k|\mathcal{F}_{k-1}^X\right] = \mathbb{E}[X_k] = 0$$

Definition 2.6. (scaled symmetric random walk) Fixed a positive integer n, define the scaled symmetric random walk,

$$W_t^{(n)} = \frac{1}{\sqrt{n}} M_{nt}$$

Provided nt is an integer. and $nt \notin \mathbb{N}$, define $W_t^{(n)}$ by linear interpolation, i.e,

$$W_t^{(n)} = ([nt] + 1 - nt) W^{(n)}_{\frac{[nt]}{n}} + (nt - [nt]) W^{(n)}_{\frac{[nt]+1}{n}}$$

Remark 2.7. *The jump size is* $\frac{1}{\sqrt{n}}$*, and the frequency is* $\frac{1}{n}$*, and we could connect all the points with straight lines*

Theorem 2.8. *The scaled symmetric random walk has independent increments*

Proof: If $0 = t_0 < t_1 < t_2 < t_3 ... < t_m = t$ satisfy $nt_i \in \mathbb{N}$ for all i, then

$$W_{t_1}^{(n)}, W_{t_2}^{(n)} - W_{t_1}^{(n)}, W_{t_3}^{(n)} - W_{t_2}^{(n)}, ..., W_{t_m}^{(n)} - W_{t_{m-1}}^{(n)}, ...$$

are independent.

Theorem 2.9. (Functional Central Limit theorem) Fixed $t \geq 0$. As $n \to \infty$, the distribution of scalar symmetric random walk $(W_t^{(n)})$ evaluated at time t converges to $\mathcal{N}(0, t)$ in the distribution

Remark 2.10. *The Functional Central Limit theorem would be met twice in Econometrics II, one is in Prof. Yu Jun's part, and the other in Associate Prof. Anthony Tay's part. So you guys are supposed to be able to use this theorem* 100% *correctly.*

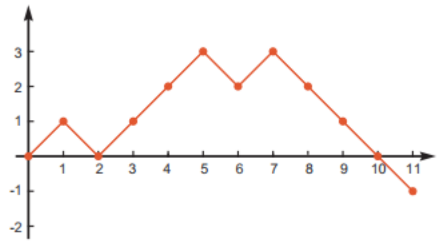


Figure 1: When n=1

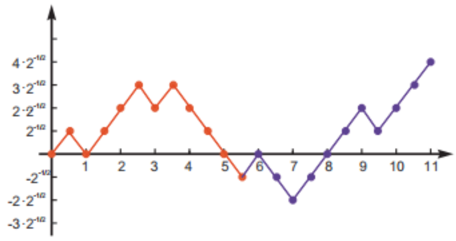


Figure 2: When n=2

2.2 Brownian Motions

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.

Definition 2.11. (Brownian Motions) A stochastic process $W = (W_t)_{t \geq 0}$ is called a standard Brownian Motion if

- $W_0 = 0, \mathbb{P} - a.s.$
- (W_t) has independent increments, i.e., for $0 \leq t_1 \leq t_2 \leq ... \leq t_m$

  $$W_{t_1}, W_{t_2} - W_{t_1}, W_{t_3} - W_{t_2}, ..., W_{t_m} - W_{t_{m-1}}, ...$$

  are independent
- For $0 \leq s < t$, $W_t - W_s \sim \mathcal{N}(0, t-s)$

Remark 2.12. *Difference between Brownian Motion (W_t) and scaled symmetric random walk $W_t^{(n)}$.*

- *The scaled random walk has a natural time step $\frac{1}{n}$ and is linear between these time steps.*
- *The scaled random walk $W_t^{(n)}$ is only approximated normal for each t, but Brownian motion is exactly normal*
- *$lim_{n\to\infty} W_t^{(n)}$ could be treated as Brownian Motions*

Theorem 2.13. *For $0 \leq s \leq t$, the covariance of W_s and W_t is as. Explicitly,*

$$\mathbb{E}[W_s W_t] = s \wedge t$$

Proof: Since $\mathbb{E}[W_s] = \mathbb{E}[W_t] = 0$, the covariance of W_s and W_t is given by,

$$\mathbb{E}[W_s W_t] = \mathbb{E}[W_s(W_t - W_s + W_s)] = \mathbb{E}[W_s(W_t - W_s)] + \mathbb{E}\left[W_s^2\right]$$

$$= \mathbb{E}[W_s][(W_t - W_s)] + \mathbb{E}\left[W_s^2\right] = 0 + s + s$$

Theorem 2.14. *Brownian Motion is a Martingale*

Proof: For $0 \leq s < t$,

$$\mathbb{E}[W_t|\mathcal{F}_s] = \mathbb{E}[W_t - W_s|\mathcal{F}_s] - \mathbb{E}[W_s|\mathcal{F}_s] = \mathbb{E}[W_t - W_s] + W_s = 0 + W_s = W_s$$

2.3 Quadratic Variation of Brownian Motions

Theorem 2.15. (Quadratic Variation) The quadratic variation of the standard Brownian motion is given by

$$\langle W \rangle_t = t, \mathbb{P} - a.s.$$

for all $t \geq 0$

Proof: (Claim: $W_t^w - t$ is martingale by Doob-Meryer Decompostion) For $0 \leq s \leq t$

$$\mathbb{E}\left[W_t^2 - t|\mathcal{F}_s\right] = \mathbb{E}\left[(W_t - W_s + W_s)^2|\mathcal{F}_s\right] - t$$

$$= \mathbb{E}\left[(W_t - W_s)^2 + 2W_s(W_t - W_s) + W_s^2|\mathcal{F}_s\right] - t$$

$$= \mathbb{E}\left[(W_t - W_s)^2|\mathcal{F}_s\right] + 2\mathbb{E}[W_s(W_t - W_s)|\mathcal{F}_s] + \mathbb{E}\left[W_s^2|\mathcal{F}_s|\mathcal{F}_s\right] - t$$

$$= \mathbb{E}\left[(W_t - W_s)^2\right] + 2\mathbb{E}[W_s(W_t - W_s)] + \mathbb{E}\left[W_s^2|\mathcal{F}_s\right] - t$$

$$= t - s + 0 + W_s^2 - t = W_s^2 - s$$

Due to the Doob-Meyer decomposition, we have $\langle M \rangle_t = t$

Remark 2.16. *Let* $\Pi = \{t_0, t_{1,\dots}\}$ *be a partition of* $[0, t]$. *Then*

$$lim_{\|\Pi\| \to 0} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})^2 = t$$

$$lim_{\|\Pi\| \to 0} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})(t_{i+1} - t_t) = 0$$

$$lim_{\|\Pi\| \to 0} \sum_{i=0}^{n-1} (t_{i+1} - t_i)^2 = 0$$

Formally, we could write it as,

$$dW_t \cdot dW_t = dt$$

$$dW_t \cdot dt = 0$$

$$dt \cdot dt = 0$$

2.4 Exponential Martingales

Suppose the interest rate $r = 0$

$$u_n = 1 + \frac{\sigma}{\sqrt{n}}, d_n = 1 - \frac{\sigma}{\sqrt{n}}$$

where σ is a positive constant (called volatility) Assume that $nt \in \mathbb{N}$, that means that up to time t, it is an nt-period model. The risk-neutral probability measure in the one-period model is given by

$$p = \frac{1 - d_n}{u_n - d_n} = \frac{1}{2}$$

$$q = \frac{u_n - 1}{u_n - d_n} = \frac{1}{2}$$

Just like to toss a fair coin. This implies that we might regard the nt-period model as tossing a fair coin nt times Suppose H_{nt} = the number of heads in the first nt coin tosses T_{nt} = the number of tails in the first nt coin tosses M_{nt} = the position of the 1-dimensional random walk. And we have

$$\begin{cases} H_{nt} + T_{nt} = nt \\ H_{nt} - T_{nt} = M_{nt} \end{cases}$$

Thus

$$\begin{cases} H_{nt} = \frac{nt+M_{nt}}{2} \\ T_{nt} = \frac{nt-M_{nt}}{2} \end{cases}$$

This implies that the stock price at time t is given by

$$S_t^n = S_0 u_n^{H_{nt}} d_n^{T_{nt}} = S_0 (1 + \frac{\sigma}{\sqrt{n}})^{\frac{nt+M_{nt}}{2}} (1 - \frac{\sigma}{\sqrt{n}})^{\frac{nt-M_{nt}}{2}}$$

Theorem 2.17. *As $n \to \infty$, the distribution of S_t^n converges to the distribution of*

$$S_t = S_0 exp(\sigma W_t - \frac{\sigma^2 t}{2})$$

where $W_t \sim \mathcal{N}(0, t)$

Proof is easy, which uses the taylor expansion rule, and it is left as the exercise.

Definition 2.18. *Let (W_t) be a Brownian Motion with filtration $(\mathcal{F}_t), \sigma \in \mathbb{R}$. The exponential martingale corresponding to σ is defined as*

$$Z_t = exp(\sigma W_t - \frac{\sigma^2 t}{2})$$

Theorem 2.19. $(Z_t, \mathcal{F}_\sqcup)_{t \geq 0}$ *is a martingale*

Proof is omitted here which exactly follows the step of previous techniques.

3 Stochastic Integrals

3.1 Construction of the Stochastic Integrals with Respect to Martingale

Let $I = [0, \infty), M \in \mathcal{M}^{2,loc}$. And the primary goal is to define an integral $\int H_s dM_s$ for a martingale M and some suitable process H. The first thought is to borrow the definition of Riemann-Stieltjes integral, which is as,

$$\int_0^t f(s) d\alpha_s = lim_{n\to\infty} \sum_{i=1}^{\infty} f(t_i^*)(\alpha(t_{i+1}) - \alpha(t_i))$$

where f is continuous, α is of bounded variation, and $t_i^* \in [t_i, t_{i+1}]$ But the problem is whether $\int H_s dM_s$ could be defined in the similar way. The answer is NO! Because by the previous results Remark 1.27, if M is a continuous non-constant local martingale, then M is not of bounded variation,and Riemann-Stieltjes integral fails here. Why? Because

$$lim_{n\to\infty} \sum_{i=1}^{n} H_{t_i}(M_{t_{i+1}} - M_{t_i}) \neq lim_{n\to\infty} \sum_{i=1}^{n} H_{t_{i+1}}(M_{t_{i+1}} - M_{t_i})$$

which contradicts the definition of Riemann-Stieltjes integral. Another question is which kind of definition is more appropriate? Using the left end point H_{t_i}, the right end point $H_{t_{i+1}}$, and the middle is $H_{\frac{t_i+t_{i+1}}{2}}$. And this brings three kinds of definition for stochastic integral

- H_{t_i}(Left-End Point): *Ito Integral*
- $H_{\frac{t_i+t_{i+1}}{2}}$(Middle Point): *Fisk-Stratonovich Integral*
- $H_{t_{i+1}}$(Right-End Point): *Backward Ito Integral*

From the perspective of Macroeconomics and Econometrics, the definition of Ito integral is much more useful, since it is feasible to know the current price from the future.

3.2 Simple Process

Remark 3.1. *Define ε^b= the collection of all bounded previsible simple processes, i.e, all processes H of the form,*

$$H_t(\omega) = \sum_{i=0}^{n-1} h^i(\omega) \mathbb{I}_{(t_i, t_{i+1}]}(t)$$

for $0 \leq t \leq \infty$

Theorem 3.2. *For $H \in \varepsilon^b$, $H \cdot M \in \mathcal{M}_0^{2,c}$. Moreover, if M is continuous, $H \cdot M$ is continuous, i.e., $H \cdot M \in \mathcal{M}_0^2$, and*

$$\mathbb{E}\left[(H \cdot M)_\infty^2\right] = \mathbb{E}\left[\int_0^\infty H_s^2 d\langle M \rangle_s\right]$$

Proof: Let $s \leq t$. If $s = t_k$, $t = t_l$ with k<l, then,

$$\mathbb{E}\left[(H\cdot M)_t - (H\cdot M)_s|\mathcal{F}_s\right] = \sum_{i=k}^{l-1}\mathbb{E}\left[h^i(M_{t_{i+1}} - M_{t_i})|\mathcal{F}_{t_k}\right]$$

$$= \sum_{i=k}^{l-1}\mathbb{E}\left[\mathbb{E}\left[h^i(M_{t_{i+1}} - M_{t_i})|\mathcal{F}_{t_i}\right]|\mathcal{F}_{t_k}\right]$$

Since h^i is $\mathcal{F}_{t_i}-$measurable and M is martingale, we have,

$$\mathbb{E}\left[h^i(M_{t_{i+1}} - M_{t_i})|\mathcal{F}_{t_i}\right] = h^i\mathbb{E}\left[(M_{t_{i+1}} - M_{t_i})|\mathcal{F}_{t_i}\right]$$

Hence,

$$\mathbb{E}\left[(H\cdot M)_t - (H\cdot M)_s|\mathcal{F}_s\right] = 0$$

Here to simplify our discussion, we only let $s = t_k, t = t_l$. Moreover, since M is a martingale,

$$\mathbb{E}\left[(H\cdot M)_\infty^2\right] = \mathbb{E}\left[\sum_{i=k}^{l-1}(h^i)^2((M_{t_{i+1}} - M_{t_i})^2\right]$$

$$= \sum_{i=k}^{l-1}\mathbb{E}\left[(h^i)^2\mathbb{E}\left[(M_{t_{i+1}} - M_{t_i})^2|\mathcal{F}_{t_i}\right]\right]$$

Since M and $M^2 - \langle M\rangle$ are martingales,

$$\mathbb{E}\left[(M_{t_{i+1}} - M_{t_i})^2|\mathcal{F}_{t_i}\right] = \mathbb{E}\left[M_{t_{i+1}}^2 - M_{t_i}^2|\mathcal{F}_{t_i}\right] = \mathbb{E}\left[\langle M_{t_{i+1}}\rangle - \langle M_{t_i}\rangle|\mathcal{F}_{t_i}\right]$$

Thus,

$$\mathbb{E}\left[(H\cdot M)_\infty^2\right] = \mathbb{E}\left\{\sum_{i=0}^{n-1}\mathbb{E}\left[(h^i)^2\langle M_{t_{i+1}}\rangle - \langle M_{t_i}\rangle|\mathcal{F}_{t_i}\right]\right\}$$

$$= \mathbb{E}\left\{\mathbb{E}\left[\sum_{i=0}^{n-1}(h^i)^2\langle M_{t_{i+1}}\rangle - \langle M_{t_i}\rangle|\mathcal{F}_{t_i}\right]\right\}$$

$$= \mathbb{E}\left[\int_0^\infty H_s^2 d\langle M\rangle_s\right]$$

Remark 3.3. *For $a < b$, we denote that,*

$$\int_a^b H_s dM_s = (H\cdot M)_b - (H\cdot M)_a$$

Theorem 3.4. *Let $H \in \varepsilon^b$ and let W be the standard Brownian motion, then*

$$\mathbb{E}\left[\int_a^b H_s dW_s\right] = 0$$

$$\mathbb{E}\left[(\int_a^b H_s dW_s)^2\right] = \mathbb{E}\left[\int_a^b H_s^2 ds\right]$$

Theorem 3.5. *Let $H^1, H^2 \in \varepsilon^b$ and $c_1, c_2 \in \mathbb{R}$, then $c_1H^1 + c_2H^2 \in \varepsilon^b$, and*

$$((c_1H^1 + c_2H^2)\cdot M)_\infty = (c_1H^1\cdot M)_\infty + (c_2H^2\cdot M)_\infty$$

3.3 Square-integrable Processes

Theorem 3.6. *If M is any continuous martingale, H satisfies*

$$\mathbb{E}\left[\int_0^T H_s^2 d\left\langle M\right\rangle_s\right] < \infty$$

, for each $T > 0$, then there exists a sequence of simple process $H^{(n)}$ such that

$$\sup_{T>0} \lim_{n\to\infty} \mathbb{E}\left[\int_0^T |H_s^{(n)} - H_s|^2 \left\langle M\right\rangle_s\right] = 0$$

Remark 3.7. *This theorem is for any process H, which satisfies $\mathbb{E}\left[\int_0^T H_s^2 d\left\langle M\right\rangle_s\right] < \infty$, and we could find one simple process sequence $H^{(n)}$ to approximate H.*

Definition 3.8. *The stochastic integral of H with respect to the martingale M is defined by*

$$\int_0^T H_s dM_s \doteq \lim_{n\to\infty} \int_0^T H_s^{(n)} dM_s = \lim_{n\to\infty} \sum_{i=1}^{n} H_i^{(n)}(M_{i+1} - M_i)$$

in the sense of L^2-sense, where $H^{(n)}$ is a sequence of simple process satisfying $sup_{t>0} lim_{n\to\infty} \mathbb{E}\left[\int_0^T |H_s^{(n)} - H_s|^2 \left\langle M\right\rangle_s\right] = 0$

Remark 3.9.

- *The definition is well-defined, suppose there exists another sequence of simple processes $K^{(n)}$ converging to H in the sense of*

  $$sup_{t>0} \lim_{n\to\infty} \mathbb{E}\left[\int_0^T |H_s^{(n)} - H_s|^2 \left\langle M\right\rangle_s\right] = 0$$

  Then the sequence $Z^{(n)}$ with $Z^{(2n-1)} = H^{(n)}$, and $Z^{(2n)} = K^{(n)}$ is also convergent to H in the sense of $\sup_{t>0} \lim_{n\to\infty} \mathbb{E}\left[\int_0^T |H_s^{(n)} - H_s|^2 \left\langle M\right\rangle_s\right] = 0$. Thus we have,

  $$\lim_{n\to\infty} \int_0^T Z_s^{(n)} dM_s$$

  converges in the sense of L^2-sense, and this means that

  $$\lim_{n\to\infty} \int_0^T H_s^{(n)} dM_s = \lim_{n\to\infty} \int_0^T Z_s^{(n)} dM_s = \lim_{n\to\infty} \int_0^T K_s^{(n)} dM_s$$

Theorem 3.10. *Let H, K are square-integrable, M is martingale. Then*

- *$(\int_0^t H_s dM_s)_{0\leq t\leq T}$ is square-integrable martingale*
- *$\int_0^T (\alpha H_s + \beta K_s) dM_s = \alpha \int_0^T H_s dM_s + \beta \int_0^T K_s dM_s$*
- *$\mathbb{E}\left[(\int_0^T H_s dM_s)^2\right] = \mathbb{E}\left[H_s^2 d\left\langle M\right\rangle_s\right]$*
- *$\mathbb{E}\left[(\int_0^T H_s dM_s)^2 |\mathcal{F}_s\right] = \mathbb{E}\left[H_s^2 d\left\langle M\right\rangle_s |\mathcal{F}_s\right]$*

- $\left\langle \int_0^{\cdot} H_s dM_s \right\rangle_t = \int_0^t H_s^2 d\langle M\rangle_u$

Corollary 3.11. *If H is square-integrable, $M = W$ =Brownian motion, then,*

$$\mathbb{E}\left[\int_0^t H_u dW_u | \mathcal{F}_s\right] = 0$$

$$\mathbb{E}\left[(\int_0^t H_u dW_u)^2 | \mathcal{F}_s\right] = \mathbb{E}\left[\int_s^t H_u^2 du | \mathcal{F}_s\right] = \int_s^t \mathbb{E}\left[H_u^2 | \mathcal{F}_s\right]$$

Theorem 3.12. (Kunita-Watanabe) If M, N are martingales, H, and K are martingales, then,

$$\int_0^t |H_u K_u| d\langle M, N\rangle_u \leq (\int_0^t H_u^2 d\langle M\rangle_u)^{\frac{1}{2}} (\int_0^t K_u^2 d\langle M\rangle_u)^{\frac{1}{2}}$$

3.4 Ito Lemma

Exercise 3.13. *In calculus, we see that if $F, G \in C^1$, by chain rule we have*

$$(F \circ G)^{'} = (F^{'} \circ G) \cdot G^{'}$$

or in differential form

$$\frac{d}{dt}(F(G(t))) = F^{'}(G(t)) \cdot \frac{dG(t)}{dt} = F^{'}(G(t)) \cdot G^{'}(t)$$

This implies that

$$F(G(t)) - F(G(0)) = \int_0^t F^{'}(G(s)) G^{'}(s) ds = \int_0^t F^{'}(G(s)) dG(s)$$

But for stochastic calculus, the last equation would not hold for two reasons, the first is that the differentials for martingales and semimartingales are not defined, and the second is that there exists great differences between Ito integral and the Riemann integral.

Theorem 3.14. (one-dimensional Ito formula, continuous form) Let $f : \mathbb{R} \to \mathbb{R}$ be a C^2- function and let $X = (X_t, \mathcal{F}_t)$ be a continuous semimartingale with the decomposition,

$$X_t = X_0 + M_t + A_t$$

where M is a local martingale and A is of bounded variation. Then

$$f(X_t) = f(X_0) + \int_0^t f^{'}(X_s) dX_s + \frac{1}{2}\int_0^t f^{''}(X_s) d[X, X]_s$$

$$= f(X_0) + \int_0^t f^{'}(X_s) dM_s + \int_0^t f^{'}(X_s) dA_s + \frac{1}{2}\int_0^t f^{''}(X_s) d\langle M\rangle_s$$

Remark 3.15. *In the differential form*

$$df(X_t) = f^{'}(X_s) dX_s + \frac{1}{2} f^{''}(X_s) d[X, X]_s$$

$$= f^{'}(X_s) dM_s + f^{'}(X_s) dA_s + \frac{1}{2} f^{''}(X_s) d\langle M\rangle_s$$

Note that $f^{'}(X_s) dM_s$ is a local martingale, $f^{'}(X_s) dA_s + \frac{1}{2} f^{''}(X_s) d\langle M\rangle_s$ is of bounded variation. This means that if X is (continuous) semimartingale and $f \in C^2$, then $f(X)$ is again a semimartingale

Proof: this proof is only basic idea, not that rigorous. By Taylor expansion,

$$f(X_{t_{i+1}\wedge t}) - f(X_{t_i}) = f^{'}(X_{t_i})\Delta_i X + \frac{1}{2}f^{''}(X_{t_i})(\Delta_t X)^2 + R_i$$

where $\Delta_i X = X_{t_{i+1}\wedge t} - X_{t_i \wedge t}$ and R_i is the error term. Summarize the above term, we get

$$f(X_t) - f(X_0) = \sum f^{'}(X_{t_t})\Delta_i X + \frac{1}{2}\sum f^{''}(X_{t_t})(\Delta_i X)^2 + \sum R_i$$

Due to the definition of stochastic integral and Riemann-Stieltjes integral, we have

$$\sum f^{'}(X_{t_t})\Delta_i X \to \int_0^t f^{'}(X_{t_t})\Delta_i X$$

$$\frac{1}{2}\sum f^{''}(X_{t_t})(\Delta_i X)^2 \to \frac{1}{2}\int_0^t f^{''}(X_{t_t})(\Delta_i X)^2$$

as $n \to \infty$ and

$$|\sum R_i| \leq \frac{1}{2}\sum |f^{''}(X_{t_i}) - f^{''}(\widetilde{X_{t_i}})|(\Delta_t X)^2 \leq \frac{\epsilon}{2}\sum (\Delta_i X)^2 \to 0$$

for large n enough, where $\widetilde{X_{t_i}}$ is between X_{t_i} and $X_{t_{i+1}\wedge t}$

Remark 3.16. *How to calculate $[X, X]$? If X is continuous semimartingale, then in notation*

$$(dX_t)^2 = d\,[X, X]_t$$

Thus, the previous one formula is converted into the following one as,

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

Moreover, if $X = W$ is standard Brownian motion, then we have

$$(dt)^2 = dt \cdot dW_t = dW_t \cdot dt = 0$$

$$(dW_t)^2 = d\,\langle W \rangle_t = dt$$

Exercise 3.17. *Consider $X = W$ = standard Brownian motion and*

$$f(x) = x^2, f^{'}(x) = 2x, f^{''}(x) = 2$$

thus

$$W_t^2 = W_0^2 + \int_0^t f^{'}(W_s)dW_s + \int_0^t \frac{1}{2}f^{''}(W_s)d\,\langle W_s \rangle_s$$

$$= 2\int_0^t W_s)dW_s + \frac{1}{2}\int_0^t f^{''}(W_s)(ds)^2 = 2\int_0^t W_s dW_s + t$$

i.e.,

$$\int_0^t W_s dW_s = \frac{1}{2}(W_t^2 - t)$$

Theorem 3.18. (multi-dimensional Ito formula, continuous local martingale) Let $X = (X^1, X^2, ..., X^n)$ be a vector of local martingales in $\mathcal{M}^{c,loc}$. Let $f : [0,\infty) \times \mathbb{R}^n \to \mathbb{R}$ be a $C^{1,2}-$function. Then

$$f(t, X_t) = f(0, X_0) + \int_0^t \frac{\partial}{\partial x} f(s, X_s) ds + \sum_{i=1}^n \int_0^t \frac{\partial}{\partial x} f(s, X_s) dX_s^i$$

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

for all t

Theorem 3.19. (Ito formula, general form) Let $X = (X^1, X^2, ...X^n)$ be an n-dimensional semimartingale with decompostion

$$X_t^i = X_0^i + M_t^i + A_t^i$$

for $i \leq i \leq n$, where M^i is a local martingale and A^i is of bounded variation. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a C^2 function, then $f(X)$ is a semimartingale and

$$f(X_t) = f(X_0) + \sum_{i=1}^n \int_0^t \frac{\partial}{\partial x_i} f(X_u) dX_u^i$$

$$+ \frac{1}{2} \sum_{i,j=1}^n \int_0^t \frac{\partial^2}{\partial x_i \partial x_j} f(X_u) d \left\langle M^{i,c}, M^{j,c} \right\rangle_u$$

$$+ \sum_{s \leq t} \left[f(X_s) - f(X_{s-}) - \sum_{i=1}^n \frac{\partial}{\partial x_i} f(X_{s-}) \Delta X_{s-} \right]$$

Exercise 3.20. (stochastic dynamic programming, example used in Macro II) Assume $y = F(t, x)$, and $dx = g(t, x)dt + \sigma(t, x)dz$, and the differential of y is as,

$$\begin{aligned} dy &= F_t dt + F_x dx + \frac{1}{2} F_{tt} (dt)^2 + F_{tx} dt dx + \frac{1}{2} F_{xx} (dx)^2 \\ &= F_t dt + F_x dx + F_{tx} dt dx + \frac{1}{2} F_{xx} (dx)^2 \\ &= F_t dt + F_x (g dt + \sigma dz) + F_{tx} dt (g dt + \sigma dz) + \frac{1}{2} (g dt + \sigma dz)^2 \qquad (1) \\ &= F_t dt + F_x g dt + F_x \sigma dz + \frac{1}{2} \sigma^2 dt \\ &= (F_t + F_x g + \frac{1}{2} F_{xx} \sigma^2) dt + F_x \sigma dz \end{aligned}$$

And for another y, $y = F(t, x_1, x_2)$ and

$$dx_1 = g_1(t, x_1)dt + \sigma_1(t, x_1)dz_1$$

$$dx_2 = g_2(t, x_2)dt + \sigma_2(t, x_2)dz_2$$

and

$$dz_1 dz_2 = \rho dt$$

The differential of y is as

$$dy = F_t dt + F_{x_1} dx_1 + F_{x_2} dx_2$$
$$+ \frac{1}{2}(F_{tt}(dt)^2 + F_{x_1x_1}(dx_1)^2 + F_{x_2x_2}(dx_2)^2 + F_{tx_1} dt dx_1 + F_{tx_2} dt dx_2 + F_{x_1x_2} dx_1 dx_2)$$
$$= (F_t + F_{x_1} g_1 + F_{x_2} g_2 + \frac{1}{2}(F_{x_1x_1}\sigma_1^2 + F_{x_2x_2}\sigma_2^2) + F_{x_1x_2}\sigma_1\sigma_2\rho)dt + F_{x_1}\sigma_1 dz_1 + F_{x_2}\sigma_2 dz_2 \quad (2)$$

And we could now start to solve a continuous-time stochastic problem, and the optimal value function is as

$$J(t_0, x_0) = max_u \mathbb{E}_{t_0} \int_{t_0}^{T} f(t, x, u)dt + \theta(x(T), T)$$
$$s.t., dx = g(t, x, u)dt + \sigma(t, x, u)dz$$

Now let us write the value function recursively

$$J(t, x) = max_u \mathbb{E}_t \int_t^T f(s, x, u)ds + \theta(x(T), T)$$
$$= max_u \mathbb{E}_t \left\{ \int_t^{t+\Delta t} f(s, x, u)ds + (E_{t+\Delta t} \int_{t+\Delta t}^{T} f(s, x, u)ds + \theta(x(T), T)) \right\}$$
$$= max_u \mathbb{E}_t \left\{ \int_t^{t+\Delta t} f(s, x, u)ds + J(t + \Delta t, x + \Delta x) \right\}$$
$$= max_u \mathbb{E}_t \left\{ f(\tilde{t}, x(\tilde{t}), u(\tilde{t}))\Delta t + J(t + \Delta t, x + \Delta x) \right\}$$
$$= max_u \mathbb{E}_t \left\{ f(\tilde{t}, x(\tilde{t}), u(\tilde{t}))\Delta t + J(t, x) + J_t \Delta t + J_x \Delta x + \frac{1}{2} J_{xx} (\Delta x)^2 \right\}$$
$$= max_u \mathbb{E}_t \left\{ f(\tilde{t}, x(\tilde{t}), u(\tilde{t}))\Delta t + J(t, x) + J_t \Delta t + J_x g \Delta t + J_x \sigma \Delta z + \frac{1}{2} J_{xx} \sigma^2 \Delta t \right\} \quad (3)$$

Thus we have

$$0 = max_u \mathbb{E}_t (f(\tilde{t}, x(\tilde{t}), u(\tilde{t}))\Delta t + J_t \Delta t + J_x g \Delta t + J_x \sigma \Delta z + \frac{1}{2} J_{xx} \sigma^2 \Delta t)$$

First take expectation operator. Note that $\mathbb{E}_t \Delta z = 0$. Then dividing Δt on both sides and letting $\Delta t \to 0$. Then divide Δt on both sides and letting $\Delta t \to 0$, we have

$$0 = max_u \mathbb{E}_t (f(t, x, u) + J_t + J_x g + \frac{1}{2} J_{xx} \sigma^2$$

And this equation is called the HJB equation, which would be widely used. Using first order condition, we can solve optimal u^* as a function of J_x and J_{xx}. Then plugging the expression of optimal u^* into the HJB equation, we could solve a partial differential equation of $J(t, x)$,

$$-J_t = f(t, x, u^*) + J_x g(t, x, u^*) + \frac{1}{2}(\sigma(t, x, u^*))^2$$

This partial differential equation plus its boundary condition

$$J(T, x(T)) = \theta(x(T), T)$$

should have a solution of $J(t, x)$.

Chapter III: Ordinary Differential Equation & Stochastic Differential Equation

Liu Yanbo

May 24, 2018

Abstract

Instead of showing rigorous theory on ordinary differential equation, I would just offer basic solution techniques for first-order ordinary differential equation. Moreover, the combination of ordinary differential equation and stochastic calculus would be the natural extension. This chapter is based on the following materials.

- Blanchard, Paul, Robert L. Devaney, and R. Glen. "Hall, and Jong-Eao Lee." Differential Equations: A Contemporary Approach, Thomson Learning (2007).
- Øksendal, Bernt. "Stochastic differential equations." Stochastic differential equations. Springer Berlin Heidelberg, 2003. 65-84.
- Lecture Notes of "Ordinary Differential Equation", National Chiao Tung University, http://ocw.nctu.edu.tw/
- Lecture Notes of "Introduction to Financial Mathematics II", National Chiao Tung University, http://ocw.nctu.edu.tw/

1 Ordinary Differential Equation

1.1 Separable Equation

1.1.1 Direct Transformation

Claim 1.1.

$$y^{'} = g(x)h(y)$$

which could be separated as,

$$\frac{dy}{h(y)} = g(x)dx$$

and if we take integral on both sides, and derive the following as,

$$\int \frac{dy}{h(y)} = \int g(x)dx + C$$

Example 1.2.

$$y^{'} = \frac{y}{1+x}$$

and it could be separated as,

$$\frac{dy}{dx} = \frac{y}{1+x}$$

and take integral on both sides, and we have

$$\int \frac{dy}{y} = \int dx(1+x)$$

and we generate the following form solution as,

$$ln|x| = ln|1+x| + C$$

where C is one constant, and the general solution is given as,

$$y = e^C(1+x)$$

Claim 1.3.

$$M_1(x) \cdot M_2(y)dx + N_1(x) \cdot N_2(y)dy = 0$$

and it could separated as,

$$\frac{M_1(x)}{N_1(x)}dx + \frac{N_2(y)}{M_2(y)}dy = 0$$

and take integral for both sides and we derive the following as,

$$\int \frac{M_1(x)}{N_1(x)}dx + \int \frac{N_2(x)}{M_2(y)} = C$$

Example 1.4.

$$\frac{dy}{dx} = \frac{xy + 3x - y - 3}{xy - 2x + 4y - 8}$$

and separate two sides as,

$$(xy + 3x - y - 3)dx = (xy - 2x + 4y - 8)dy$$

$$(x-1)(y+3)dx = (x+4)(y-2)dy$$

$$(\frac{x-1}{x+4})dx = (\frac{y-2}{y+3})dy$$

$$(1 - \frac{5}{x+4})dx = (1 - \frac{5}{x+3})dy$$

$$\int (1 - \frac{5}{x+4})dx = \int (1 - \frac{5}{x+3})dy$$

and the general solution is as,

$$x - 5\ln x + 4 = y - 5\ln y + 3 + C$$

1.1.2 Indirect Transformation

Definition 1.5. *If the ODE could be written as* $y^{'} = f(x,y)$*, and this equation is called first-order homogeneous ODE*

Remark 1.6. *Solution procedure could be generalized as,*

- *transform* $f(x,y)$ *into* $f(\frac{y}{x})$
- *let* $\frac{y}{x} = u$*, and* $y = xu, dy = xdu + udx$
- *put the above expressions back into the ODE, and we derive this as,*

$$\frac{dy}{dx} = \frac{xdu + udx}{dx} = f(u)$$

which as

$$x\frac{du}{dx} + u = f(u)$$

and we have

$$\int \frac{du}{f(u) - u} = \int \frac{dx}{x} + C = \ln|x| + C$$

Example 1.7.

$$y^{'} = \frac{y}{x} + 1$$

let $u = \frac{y}{x}$*, thus* $y = xu, dy = xdu + udx$ *and we have*

$$\frac{xdu + udx}{dx} = u + 1$$

$$x\frac{du}{dx} + u = u + 1$$

$$\int du = \int \frac{1}{x}dx$$

$$u = \ln|x| + C$$

thus the general solution is,

$$\frac{y}{x} = \ln|x| + C$$

Remark 1.8. *If the first-order ODE is as* $M(x,y)dx + N(x,y)dy = 0$*, in which* $M(x,y)$ *and* $N(x,y)$ *are both m-order function, and this ODE is first-order ODE, and the solution procedure could be shown as,*

- *transform into* $M(\frac{y}{x})dx + N(\frac{y}{x})dy = 0$
- *let* $\frac{y}{x} = u$*, and* $dy = udx + xdu$
- *put the above expression back into ODE, which is as* $M(u)dx + N(u)(udx + xdu) = 0$*, and organize the equation, we derive the following expression as,*

$$\int \frac{dx}{x} + \int \frac{N(u)}{uN(u) + M(u)}du = C$$

Example 1.9.

$$y^{'} = \frac{x-y}{x+y}$$

$$(x-y)dx = (x+y)dy$$

$$(1-\frac{y}{x})dx = (1+\frac{y}{x})dy$$

let $u = \frac{y}{x}$, $y = xy, dy = xdu + udx$, *and we have*

$$(1-u)dx = (1+u)(xdu + udx)$$

$$\frac{1-2u+u^2}{1+u} = x\frac{du}{dx}$$

$$\int -\frac{1}{x}dx = \int \frac{1+u}{u^2+2u-1}du$$

$$-\ln|x| + C_1 = \frac{1}{2}\ln|u^2+2u-1|$$

$$x^2(u^2+2u-1) = C$$

where $C = e^{2C_1}$

$$x^2(\frac{y^2}{x^2} + \frac{2y}{x} - 1) = C$$

and the general solution is

$$y^2 + 2xy - x^2 = C$$

1.1.3 $(a_1x + b_1y + c_1)dx + (a_2x + b_2y + c_2)dy = 0$

Claim 1.10. *If* $\frac{a_1}{a_2} \neq \frac{b_1}{b_2}$ *then the solution could be derived as,*

- *find the intersection point for* $a_1x + b_1y + c_1 = 0$ *and* $a_2x + b_2y + c_2 = 0$
- *let* $x = u + \alpha \Rightarrow dx = du$, *and* $y = v + \beta \Rightarrow dy = dv$ *put the above expressions in to the original ODE, and get the first-order homogeneous ODE as,*

$$(a_1u + b_1v)du + (a_2u + b_2v)dv = 0$$

- *solve the ODE using the method developed for first-order homogeneous ODE.*

Claim 1.11. $\frac{a_1}{a_2} = \frac{b_1}{b_2} \neq \frac{c_1}{c_2}$ *And the solution is constructed as,*

- *Let* $a_2x + b_2y = t$, *and then* $a_1x + b_1y = mt$, *and* $dy = \frac{dt - a_2dx}{b_2}$
- *put the above expressions into the original ODE, and we have*

$$(mt + c_1)dx + (t + c_2)\frac{dt - a_2dx}{b_2} = 0$$

which is as,

$$\int \frac{t + c_2}{a_2t + a_2c_2 - b_2mt - b_2c_1}dt = \int dx$$

1.1.4 $(y^{'} = f(ax + by + c))$

Claim 1.12. *The solution for TYPE III,* $(y^{'} = f(ax+by+c))$ *could be shown as,*

- *let* $t = ax + by + c$*, and* $dy = \frac{dt - adx}{b}$
- *put the above expression back into the original ODE, and we have,*

$$\frac{dy}{dx} = \frac{(\frac{dt - adx}{b})}{dx} = f(t)$$

which is as,

$$\int \frac{dt}{bf(t) + a} = \int dx + c = x + c$$

Example 1.13. *Suppose the ODE is as,* $y^{'} = tan^2(x + y)$ *and let* $u = x + y$*,* $du = dx + dy$

$$\frac{du}{dx} = 1 + tan^2 u = sec^2 u$$
$$(cos^2 u)du = dx$$
$$\int \frac{1 + cos2u}{2} du = \int dx$$
$$\frac{u}{2} + \frac{1}{4} sin(2u) = x + C$$

and the general solution is

$$\frac{x + y}{2} + \frac{1}{4} sin2(x + y) = X + C$$

1.2 Exact

1.2.1 Exact Equation

Definition 1.14. *If the first-orde ODE is like* $M(x, y)dx + N(x, y)dy = 0$*, and if there exists one function* $\theta(x, y)$ *satisfies* $M(x, y)dx + N(x, y)dy = 0$ *and we call* $M(x, y)dx + N(x, y)dy = 0$ *as the first-order exact ODE And the solution techniques are given as, If* $M(x, y)dx + N(x, y)dy = 0$ *is exact, then according to the definition, we have,*

$$d\theta = \frac{\partial \theta}{\partial x} dx + \frac{\partial \theta}{\partial y} dy = M(x, y)dx + N(x, y)dy = 0$$

, Then ODE is solved as $\theta(x, y) = c \Rightarrow d\theta = \frac{\partial \theta}{\partial x} dx + \frac{\partial \theta}{\partial y} dy = 0$*, so* $\theta(x, y) = c$ *is the general solution, and* θ *satisfies*

$$\frac{\partial \theta}{\partial x} = M(x, y) \Rightarrow \theta(x, y) = \int^x M(x, y)dx + g(y)$$
$$\frac{\partial \theta}{\partial y} = N(x, y) \Rightarrow \theta(x, y) = \int^y N(x, y)dy + h(x)$$

and we could compare the above two equations, and get $g(y)$*and* $h(x)$*, and then get* $\theta(x, y)$

Exercise 1.15.

$$(siny - ysinx)dx + (cosx + xcosy - y)dy = 0$$

let $M = siny - ysinx$*,* $N = cosx + xcosy - y$

$$\Rightarrow \frac{\partial M}{\partial y} = cosy - sinx, \frac{\partial N}{\partial x} = -sinx + cosy$$

And this ODE is exact, and solve $\theta(x,y) = C^*$*, and* $\theta(x,y)$ *satisfies that*

$$\frac{\partial \theta}{\partial x} = M = siny - ysinx$$

and

$$\frac{\partial \theta}{\partial y} = N = cosx - xcosy - y$$

Take integral on both sides, and we have

$$\theta(x,y) = xsiny + ycosx + g(y)$$

$$\theta(x,y) = ycosx + xsiny - \frac{y^2}{2} + h(x)$$

and compare the above two equations, and find out that

$$\theta(x,y) = xsiny + ycosx - \frac{y^2}{2} + C_1$$

and put the above expression into $\theta(x,y) = C^*$*, and the ODE's general solution is as,*

$$xsiny + ycosx - \frac{y^2}{x} = C$$

where $C \equiv (C^* - C_1)$

1.2.2 Modified exact equation

Definition 1.16. *If on non-exact ODE is characterized as* $M(x,y)dx+N(x,y)dy = 0$*, and if there exists one function* $I(x,y)$ *could transform the original ODE back into the exact form, which is,*

$$I(x,y)M(x,y)dx + I(x,y)N(x,y)dy = d\theta(x,y) = 0$$

then $I(x,y)$ *is called the* Integrating factor.

Remark 1.17. (How to tell an ODE is exact or not) From defintion we know that $I(x,y)M(x,y)dx+I(x,y)N(x,y)dy = d\theta(x,y) = 0$ is exact, and it satisfies,

$$\frac{\partial}{\partial y}\left[I(x,y)M(x,y)\right] = \frac{\partial}{\partial x}\left[I(x,y)N(x,y)\right]$$

and the above equation could be expanded as,

$$N(x,y)\frac{\partial I}{\partial x} - M(x,y)\frac{\partial I}{\partial y} = I(\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x})$$

If $I = I(x,y)$, then there is no solution, but if $I = I(x)$ or $I = I(y)$ only, then the previous equation could be simplified as,

- If $I = I(x)$, then the above equation is as $N\frac{dI}{dx} = I(\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x})$
- If $I = I(y)$, then the above equation is as $-M\frac{dI}{dy} = I(\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x})$

And we have the characteristic function for ODE, which is as,

$$\frac{N}{dx} = \frac{-M}{dy} = \frac{I(\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x})}{dI}$$

Separate the characteristic function of ODE, we have I

- If $I = I(x)$, then by $N\frac{dI}{dx} = I(\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x})$, we have

$$\frac{dI}{I} = (\frac{\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x}}{N})dx = f(x)dx$$

$$\Rightarrow I(x) = exp(\int f(x)dx)$$

- If $I = I(y)$, then by $-M\frac{dI}{dy} = I(\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x})$, we have

$$\frac{dI}{I} = (\frac{\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x}}{-M})dy = f(y)dx$$

$$\Rightarrow I(y) = exp(\int f(y)dy)$$

Remark 1.18.
- *If ODE $M(x,y)dx + N(x,y)dy = 0$ is not exact, And if $(\frac{\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x}}{N}) = f(x)$ only, then we have*

$$I(x) = exp(\int f(x)dx)$$

- *If ODE $M(x,y)dx + N(x,y)dy = 0$ is not exact, And if $(\frac{\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x}}{-M}) = f(y)$ only, then we have*

$$I(y) = exp(\int f(y)dy)$$

- *Put I back into original ODE, and we have*

$$\Rightarrow I(x,y)M(x,y)dx + I(x,y)N(x,y)dy = d\theta(x,y)$$

is exact, and let the solution to original ODE is $\theta(x,y) = C$, and

$$\theta(x,y)$$

satisfies

$$\frac{\partial \theta}{\partial x} = I(x,y)M(x,y), \theta(x,y) = \int^x I(x,y)M(x,y)dx + g(y)$$

$$\frac{\partial \theta}{\partial y} = I(x,y)N(x,y), \theta(x,y) = \int^y I(x,y)N(x,y)dx + g(x)$$

And compare the above two equations, and we could derive $\theta(x,y)$

Example 1.19. *Suppose the ODE is as* $6xydx + (4y + 9x^2)dy = 0$ *and Let* $M = 6xy$, *and* $N = 4y + 9x^2 \Rightarrow \frac{\partial M}{\partial y} = 6x$, *and* $\frac{\partial N}{\partial x} = 18x$. *This ODE is not exact. And check whether there exists integrating factor*

$$I(\frac{\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x}}{-M}) = \frac{-12x}{-6xy} = \frac{2}{y} = f(y)$$

only Actually it exists as

$$I(y) = exp(\int f(y)dy) = y^2$$

Plug the integrating factor back into the original ODE, and we have

$$6xy^3dx + (4y^3 + 9x^2y^2)dy = 0$$

which is an exact ODE.
Then solve $\theta(x, y) = C^*$, *and* $\theta(x, y)$ *satisfies,*

$$\frac{\partial \theta}{\partial x} = 6xy^3 \& \frac{\partial \theta}{\partial y} = N = 4y^3 + 9x^2y^2$$

Take integral with respect to x *and* y *for the above two equations, and we have*

$$\theta(x, y) = 3x^2y^3 + g(y) \& \theta(x, y) = y^4 + 3x^2y^3 + h(x)$$

and compare the above two equations, and we have

$$\theta(x, y) = 3x^2y^3 + y^4 + C_1$$

And put it back into $\theta(x, y) = C^*$, *and we have the general solution as,* $3x^2y^3 + y^4 = C$, *where* $C = (C^* - C_1)$

1.2.3 Linear Differential Equation

Definition 1.20. *A* 1^{st}*-order linear ODE as,* $y'(x) + P(x)y(x) = Q(x)$

- *When* $Q(x) = 0$, *it is called homogeneous ODE*
- *When* $Q(x) \neq 0$, *it is called nonhomogeneous ODE*

Claim 1.21. *A* $1^{st}-$*order linear ODE is like* $y'(x) + P(x)y(x) = Q(x)$

$$\Rightarrow dy + [P(x)y - Q(x)]\,dx = 0$$

Let $M(x, y) = P(x)y - Q(x)$, *and* $N(x, y) = 1$
Then $\frac{\partial M}{\partial y} = P(x), \frac{\partial N}{\partial x} = 0$, $\Rightarrow$*non-exact.*
Then check if an integrating factor exists, and we observe that,

$$\frac{\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x}}{N} = P(x)$$

x *only* $\Rightarrow$*The equation has the integrating factor* $I(x) = exp(\int P(x)dx)$ *And the original ODE could be transformed as* $I(x)y'(x) + I(x)P(x)y(x) = I(x)Q(x)$, *which is exact* $\Rightarrow (Iy)' = IQ$, *and take integral to get* $Iy = \int IQdx + C$, *which is the general solution*

1.2.4 Bernoulli's Equation

Definition 1.22. $y^{'} + P(x)y = Q(x)y^n$, *which is* 1^{st}*−order non-linear ODE*

Remark 1.23. *This kind of equation is very common in engineering and physics.*

Claim 1.24. (Solution Procedure)

$$y^{'} + P(x)y = Q(x)y^n$$

$$\Rightarrow y^{-n}y^{'} + P(x)y^{1-n} = Q(x)$$

Using the approach transformation of variable, let $u = y^{1-n}$, $\frac{du}{dx} = (1-n)y^{-n}y^{'}$
And plug it back into ODE, and we have

$$\frac{du}{dx} + (1-n)P(x)u = (1-n)Q(x)$$

And let the integrating factor be defined as $I(x) = exp(\int (1-n)P(x)dx)$

$$\Rightarrow Iu = \int I(1-n)Q(x)dx + C$$

Example 1.25. *Suppose* $xy^{'} + 2y = xy^3$

$$\Rightarrow y^{'} + \frac{2}{x}y = y^3$$

$$y^{-3}y^{'} + \frac{2}{x}y^{-2} = 1$$

Let $u = y^{-2}, u^{'} = -2y^{-3}y^{'}$ *plug the above expressions back into the original ODE, and we have*

$$-\frac{u^{'}}{2} + \frac{2}{x}u = 1$$

, and we have $u^{'} - \frac{4}{x}u = -2$

$$\Rightarrow P(x) = \frac{-4}{x}, Q(x) = -2$$

And the integrating factor

$$I(x) = exp\int (-\frac{4}{x})dx = \frac{1}{x^4}$$

$$Iu = \int \frac{1}{x^4}(-2)dx = \frac{2}{3x^2} + C$$

$$\Rightarrow u = \frac{2}{3}x + Cx^4 = y^{-2}$$

which is the general solution

2 Stochastic Differential Equation

Consider a stochastic process (X_t) satisfying

$$dX_t = b(t, X_t, W_t)dt + \sigma(t, X_t, W_t)dW_t$$

Question

- Can we obtain the existence and uniqueness theorem for $dX_t = b(t, X_t, W_t)dt + \sigma(t, X_t, W_t)dW_t$? What are the properties of the solution?
- How to solve it?

2.1 Examples and some solution methods

Example 2.1.

$$dX_t = \alpha X_t dW_t + \sigma X_t dt$$

where the initial condition X_0 is given and α, σ are constant. Rewrite this equation as

$$\frac{dX_t}{X_t} = \alpha dW_t + \sigma dt$$

and we could get

$$\int_0^t \frac{dX_t}{X_t} = \int_0^t \alpha dW_t + \int_0^t \sigma du = \alpha W_t + \sigma t$$

Due the Ito Lemma, we have

$$df(t, X_t) = \frac{\partial}{\partial t} f(t, X_t)dt + \frac{\partial}{\partial x} f(t, X_t)dX_t + \frac{1}{2}\frac{\partial^2}{\partial x^2} f(t, X_t)(dX_t)^2$$

and we might have $f(t, X_t) = f(x) = \ln x$, then

$$f^{'}(x) = \frac{1}{x}, f^{''}(x) = -\frac{1}{x^2}$$

Then

$$d\ln(X_t) = \frac{1}{X_t} dX_t - \frac{1}{2}\frac{1}{X_t^2}(dX_t)^2 = \frac{1}{X_t} dX_t - \frac{\alpha^2}{2} dt$$

i.e.,

$$\ln(X_t) - \ln(X_0) - \int_0^t \frac{dX_u}{X_u} - \frac{\alpha^2 t}{2}$$

And combine $\int_0^t \frac{dX_t}{X_t} = \int_0^t \alpha dW_t + \int_0^t \sigma du = \alpha W_t + \sigma t$, and $\ln(X_t) - \ln(X_0) - \int_0^t \frac{dX_u}{X_u} - \frac{\alpha^2 t}{2}$, we get

$$\ln(\frac{X_t}{X_0}) + \frac{\alpha^2 t}{2} = \int_0^t \frac{dX_u}{X_u} = \alpha W_t + \sigma t$$

Thus, the solution to the stochastic differential equation is given as,

$$X_t = X_0 \exp(\alpha W_t + (\sigma - \frac{\alpha^2}{2})t)$$

Definition 2.2. *A stochastic process* (X_t) *of the form* $X_t = X_0 exp(\alpha W_t + \mu t)$, *is called the geometric Brownian motion.*

Remark 2.3.

- *If* (W_t) *is independent of* X_0, *then*

$$\mathbb{E}\left[X_t\right] = \mathbb{E}\left[X_0 exp(\alpha W_t + (\sigma - \frac{\alpha^2}{2})t)\right] = \mathbb{E}\left[X_0\right] e^{\sigma t}$$

- *The solution of SDE has several properties as,*
  - *If* $\sigma > \frac{\alpha^2}{2}$, *then* $X_t \to \infty$ *as* $t \to \infty$, $\mathbb{P}-a.s.$
  - *If* $\sigma < \frac{\alpha^2}{2}$, *then* $X_t \to 0$ *as* $t \to \infty$, $\mathbb{P}-a.s.$
  - *If* $\sigma = \frac{\alpha^2}{2}$, *then* X_t *will fluctuate between arbitrary large and arbitrary small values as* $t \to \infty$, $\mathbb{P}-a.s.$

Example 2.4. (Hull-White interest-rate model) Consider

$$dR_t = (a_t - b_t R_t)dt + \sigma_t dW_t$$

with $R_0 = r$, where a_t, b_t, and σ_t are deterministic function. Then

$$dR_t + b_t R_t dt = a_t dt + \sigma_t dW_t$$

which implies that

$$d(R_t exp(\int_0^t b_u du)) = a_t exp(\int_0^t b_u du)dt + \sigma_t exp(\int_0^t b_u du)dW_t$$

Thus, the solution to the original SDE is given by,

$$R_t = r exp(-\int_0^t b_u du) + \int_0^t a_s exp(-\int_0^t b_u du)ds + \int_0^t \sigma_s exp(-\int_0^t b_u du)dW_s$$

Example 2.5. *Consider the stochastic differential equation*

$$dX_t = rX_t(K - X_t)dt + \beta X_t dW_t$$

Rewrite the equation as

$$\frac{dX_t}{X_t} + rX_t dt = rK dt + \beta dW_t$$

Taking integration on both sides and we have,

$$\int_0^t \frac{dX_u}{X_u} + r\int_0^t X_u du = rKt + \beta W_t$$

Using a similar argument as in the previous example and we have

$$\int_0^t \frac{dX_u}{X_u} = \ln X_t - \ln X_0 + \frac{1}{2}\int_0^t \frac{1}{X_u^2} d\left\langle X\right\rangle_u$$

$$= \ln(\frac{X_t}{x}) + \frac{1}{2}\int_0^t \frac{1}{X_u^2}\beta^2 X_u^2 du = \ln(\frac{X_t}{x}) + \frac{1}{2}\beta^2 t$$

Thus,

$$\ln(\frac{X_t}{x}) + r\int_0^t X_u du = \beta W_t + (rK - \frac{1}{2}\beta^2)t$$

Which implies that

$$X_t exp(r\int_0^t X_u du) = x\exp\left[\beta W_t + (rK - \frac{1}{2}\beta^2)t\right]$$

Integration with respect to t on both sides, we obtain

$$x\int_0^t \exp\left[\beta W_s + (rK - \frac{1}{2}\beta^2)s\right] ds = \int_0^t X_s exp(r\int_0^s X_u du)ds$$

$$= \int_0^t \exp(r\int_0^s X_u du) d(\int_0^s X_u du)$$

$$= \frac{1}{r}\exp(\int_0^s X_u du)|_{s=0}^t$$

$$= \frac{1}{r}\left[\exp(r\int_0^t X_u du) - 1\right]$$

Hence,

$$exp(r\int_0^t X_u du) = 1 + rx\int_0^t exp\left[\beta W_s + (rK - \frac{1}{2}\beta^2)d\right] ds$$

i.e.,

$$\int_0^t X_u du = \frac{1}{2}\ln(1 + rx\int_0^t exp\left[\beta W_s + (rK - \frac{1}{2}\beta^2)d\right] ds)$$

Taking derivative with respect to t, we see that the solution to the SDE is as,

$$X_t = \frac{1}{r}\frac{rxexp\left[\beta W_t + (rK - \frac{1}{2}\beta^t t)\right]}{1 + rx\int_0^t exp\left[\beta W_s + (rK - \frac{1}{2}\beta^2)d\right] ds}$$

$$= \frac{exp\left[\beta W_t + (rK - \frac{1}{2}\beta^t t)\right]}{x^{-1} + r\int_0^t exp\left[\beta W_s + (rK - \frac{1}{2}\beta^2)d\right] ds}$$

Example 2.6. *Solving the stochastic differential equation*

$$dX_t = a_t dt + b_t X_t dW_t$$

Rewrite it as,

$$dX_t - b_t X_t dW_t = a_t dt$$

And we are supposed to find the appropriate integrating factor to reduce the equation, such that,

$$\rho_t dX_t - b_t \rho_t X_t dW_t = a_t \rho_t dt$$

By integration by parts, we have

$$d(\rho_t X_t) = -b_t \rho_t dX_t + X_t d\rho_t + d\left\langle \rho, X \right\rangle_t$$

By the idea of integrating factor, we are to find $X_t d\rho_t = -b_t\rho_t dW_t$, thus we want to find ρ such that

$$\frac{d\rho_t}{\rho_t} = -b_t dW_t$$

Then by Ito's lemma,

$$-\int_0^t b_u dW_u = \int_0^t \frac{d\rho_u}{\rho_u} = \ln(\rho_t) - \ln(\rho_0) + \frac{1}{2}\int_0^t \frac{1}{\rho_u^2}(d\rho_u)^2$$

which is

$$-\int_0^t b_u dW_u = \ln(\rho_t) - \ln(\rho_0) + \frac{1}{2}\int_0^t b_u^2 du$$

$$\Leftrightarrow \rho_t = \rho_0 \exp(-\int_0^t b_u dW_u - \frac{1}{2}\int_0^t b_u^2 du)$$

Hence,

$$d(\rho_t X_t) = \rho_t dX_t + X_t d\rho_t + (dX_t)(d\rho_t)$$

$$= \rho_t dX_t - b_t\rho_t dW_t + (a_t dt + b_t X_t dW_t)(-b_t\rho_t X_t dW_t)$$

$$= \rho_t dX_t - b_t\rho_t X_t dW_t - b_t^2\rho_t X_t dt$$

Plug the above result into $dX_t - b_t X_t dW_t = a_t dt$, we have

$$d(\rho_t X_t) + b_t^2\rho_t X_t dt = a_t\rho_t dt$$

Using the integrating factor again, and let $G_t = exp(\int_0^t b_u^2 du)$, we get

$$d(G_t\rho_t)X_t = G_t d(\rho_t X_t) + \rho_t X_t dG_t = a_t\rho_t G_t dt$$

set

$$F(t) = \rho_t G_t = \exp(-\int_0^t b_u dW_u + \frac{1}{2}\int_0^t b_u^2 du)$$

Thus,

$$d(F_t X_t) = a_t F_t dt$$

i.e.,

$$F_t X_t - F_0 X_0 = \int_0^t a_u F_u du$$

Hence its solution is given by,

$$X_t = F_t^{-1}X_0 + F_t^{-1}\int_0^t a_u F_u du$$

$$= X_0 \exp(\int_0^t b_u dW_u - \frac{1}{2}\int_0^t b_u^2 du) + \int_0^t a_u exp(\int_0^t b_v dW_v - \frac{1}{2}\int_0^t b_v^2 dv) du$$

Remark 2.7. (How to find integrating factor) Suppose $X_t d\rho_t = -b_t\rho_t X_t dW_t$ Hence, we want to find ρ such that,

$$\frac{d\rho_t}{\rho_t} = -b_t dW_t$$

By Ito lemma,

$$-\int_0^t b_u dW_u = \int_0^t \frac{d\rho_u}{\rho_u}$$

$$= ln(\rho_t) - ln(\rho_0) + \frac{1}{2}\int_0^t \frac{1}{\rho_u^2}(d\rho_u)^2$$

$$= ln(\rho_t) - ln(\rho_0) + \frac{1}{2}\int_0^t b_u^2 du$$

Hence, we have

$$\rho_t = \rho_0 exp(-\int_0^t b_u dW_u - \frac{1}{2}\int_0^t b_u^2 du)$$

And by letting $\rho_0 = 1$, we have

$$\rho_t = exp(-\int_0^t b_u dW_u - \frac{1}{2}\int_0^t b_u^2 du)$$

Example 2.8. *Solving the stochastic differential equation*

$$LQ_t^{''} + RQ_t^{'} + \frac{1}{C}Q_t = G_t + \alpha\widehat{W_t}$$

where $\widehat{W_t}$ is the white noise. Introduce the vector

$$X_t = \begin{Bmatrix} X_t^{(1)} \\ X_t^{(2)} \end{Bmatrix} = \begin{Bmatrix} Q_t \\ Q_t^{'} \end{Bmatrix}$$

Then

$$\begin{cases} (X_t^{(1)})' = X_t^{(2)} \\ L(X_t^{(2)})' + RX_t^{(2)} + \frac{1}{C}X_t^{(1)} = G_t + \alpha\widehat{W_t} \end{cases}$$

And we might rewrite the original SDE in this form as,

$$dX_t = AX_t dt + H_t dt + KdW_t$$

Where (W_t) is a 1-dimensional Brownian motion

$$X_t = \begin{Bmatrix} X_t^{(1)} \\ X_t^{(2)} \end{Bmatrix}$$

$$A = \begin{Bmatrix} 0 & 1 \\ -\frac{1}{CL} & -\frac{R}{L} \end{Bmatrix}$$

$$H_t = \begin{Bmatrix} 0 \\ \frac{G_t}{L} \end{Bmatrix}$$

$$K = \begin{Bmatrix} 0 \\ \frac{\alpha}{L} \end{Bmatrix}$$

Thus

$$d\exp\left[\exp(-At)X_t\right] = \exp(-At)(H_t dt + KdW_t)$$

The solution is of the form

$$X_t = \exp(At)X_0 + \exp(A_t)\int_0^t \exp(-As)(H_s ds + KdW_s)$$

2.2 An Existence and Uniqueness Result

Theorem 2.9. (Existence and uniqueness theorem for stochastic differential equation) Let $T > 0$, and let $b, \sigma, (b : [0,T] \times \mathbb{R}^n \to \mathbb{R}^n, \sigma : [0,T] \times \mathbb{R}^n \to \mathbb{R}^{n\times m})$ be measurable functions satisfying

$$|b(t,x)| + |\sigma(t,x)| \leq C(1 + |x|)$$

for all $x \in \mathbb{R}^n$, $t \in [0,T]$ for some constant C, where

$$|\sigma|^2 = \sum_{i=1}^{n} \sum_{j=1}^{m} |\sigma_{ij}|^2$$

and

$$|b(t,x) - b(t,y)| + |\sigma(t,x) - \sigma(t,y)| \leq D|x - y|$$

for all $x, y \in \mathbb{R}^n, t \in [0,T]$ for some constant D. Let Z be a random variable which is independent of$\mathcal{F}_T^X$, the $\sigma-$algebra generate by $(W_s : 0 \leq s \leq T)$, and $\mathbb{E}|Z|^2 < \infty$. Then the differential equation

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t$$

for $0 \leq t \leq T$, $X_0 = Z$ has one unique t-continuous solution X_t with the properties that

- X_t is adapted to $\mathcal{F}_t^W \vee \sigma(Z)$
- $\mathbb{E}\left[\int_0^T |X_t|^2 dt\right] < \infty$

Example 2.10. *Consider the differential equation*

$$\frac{dX_t}{dt} = 3X_t^{\frac{2}{3}}$$

with $X_0 = 0$, which has more than one solution. For any $a > 0$, the function

$$X_t = \begin{cases} 0 & for \quad t \leq a \\ (t-a)^3 & for \quad t > a \end{cases}$$

is the solution of the original SDE. In this case, $b(x) = 3x^{\frac{2}{3}}$ does not satisfy the condition $|b(t,x) - b(t,y)| + |\sigma(t,x) - \sigma(t,y)| \leq D|x - y|$ at $x = 0$

2.3 Weak and Strong solutions

Definition 2.11.

- *A strong solution of the stochastic differential equation*

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t$$

on a given probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and with respect to the fixed Brownian motion W and initial value Z, is a stochastic process X with continuous sample paths and with the following properties:

  - *X is adapted to the filtration $(\mathcal{F}_t)$*
  - $\mathbb{P}[X_0 = Z] = 1$

- $\mathbb{P}\left[\int_0^t(|b_i(s,x_s)| + \sigma_{ij}^2(s,X_s))ds\right] = 1$ *for all* $1 \leq i \leq d$, $1 \leq j \leq r$, *and* $0 \leq t < \infty$
- *the integral version of the original SDE is*

$$X_t = X_0 + \int_0^t b(s,X_s)ds + \int_0^t \sigma(s,X_s)dW_s$$

- *A weak solution of* $dX_t = b(t,X_t)dt + \sigma(t,X_t)dW_t$ *is a triple* (X,W), $(\Omega,\mathcal{F},\mathbb{P})$, $(\mathcal{F}_t)$, *where*
  - $(\Omega,\mathcal{F},\mathbb{P})$ *is a probability space and* $(\mathcal{F}_t)$ *is a filtration of sub-*$\sigma-$*algebra of* $\mathcal{F}$ *satisfying the usual conditions*
  - $X = (X_t,\mathcal{F}_t)_{0\leq t<\infty}$ *is continuous, adapted* $\mathbb{R}^n-$*valued process.* $W = (W_t,\mathcal{F}_t)_{0\leq t<\infty}$ *is a standard Brownian motion, and satisfies the following conditions as,*
    * $\mathbb{P}\left[\int_0^t(|b_i(s,x_s)| + \sigma_{ij}^2(s,X_s))ds\right] = 1$ *for all* $1 \leq i \leq d$, $1 \leq j \leq r$, *and* $0 \leq t < \infty$
    * *the integral version of the original SDE is*

$$X_t = X_0 + \int_0^t b(s,X_s)ds + \int_0^t \sigma(s,X_s)dW_s$$

Remark 2.12. *There are stochastic differential equations which has no strong solutions, but still has a weak solution*

Remark 2.13. *The major difference between strong and weak conditions are that, the probability space for strong condition is given in advance, but the one for weak condition is generated by the Brownian motion used in the stochastic differential equation.*

Chapter IV: Difference Equation & Stochastic Differential Equation

Liu Yanbo

May 24, 2018

Abstract

Difference Equation could be treated as the discrete-time ordinary differential equation, and widely used in macroeconomics, finance and time series. This chapter is based on the following materials.

- Hamilton, James Douglas. Time series analysis. Vol. 2. Princeton: Princeton university press, 1994.
- Stokey, Nancy L. Recursive methods in economic dynamics. Harvard University Press, 1989.
- Lecture Notes of "Probability Theory", University of Cambridge, https://www.cl.cam.ac.uk/teaching/2003/Probability/.
- Lecture Notes of "Introduction to Financial Mathematics II", National Chiao Tung University, http://ocw.nctu.edu.tw/

1 Introduction

Many problems in Probability give rise to difference equations. Difference equations relate to differential equations as discrete mathematics relates to continuous mathematics.

As with differential equations, one can refer to the order of a difference equation and note whether it is linear or non-linear and whether it is homogeneous or inhomogeneous. The present discussion will almost exclusively be confined to linear second order difference equations both homogeneous and inhomogeneous.

2 Notation Convention

A trivial example stems from considering the sequence of odd numbers starting from 1. The associated difference equation might be specified as:

$$f(n) = f(n-1) + 2, f(1) = 1$$

In words: term n in the sequence is two more than term $n - 1$. The proviso, $f(1) = 1$, constitute an initial condition. The first term in the sequence is 1. A term like $f(n)$ so strongly suggests a continuous function that many writers prefer to use a subscript notation. The present difference equation would be presented as:

$$u_n = u_{n-1} + 2, u_1 = 1$$

This is the notation which will be used below. It is strongly implicit that n is an integer. In simple cases, a difference equation gives rise to an associated *Auxiliary Equation*. In the case of $u_n = u_{n-1}+2, u_1 = 1$, the associated auxiliary equation is as,

$$w^1 - 1 = 0$$

The highest power of the polynomial in w is 1 and, accordingly, $u_n = u_{n-1} + 2, u_1 = 1$ is said to be a first order difference equation. If the constant term 2 were absent from $u_n = u_{n-1} + 2, u_1 = 1$, the equation would be homogeneous but its presence makes it inhomogeneous. Some standard techniques for solving elementary difference equations analytically will now be presented...

3 Second Order Homogeneous Linear Difference Equation-I

To solve,

$$u_n = u_{n-1} + 2, with, u_1 = 1, u_0 = 1$$

transfer all the terms to the left-hand side:

$$u_n - u_{n-1} - u_{n-2} = 0$$

The zero on the right-hand side signifies that this is a homogeneous difference equation. Guess:

$$u_n = Aw^n$$

so

$$Aw^n - Aw^{n-1} - Aw^{n-2} = 0$$

and

$$w^2 - w - 1 = 0$$

This is the auxiliary equation associated with the difference equation. Being a quadratic, the auxiliary equation signifies that the difference equation is of second order. The two roots are readily determined:

$$w_1 = \frac{1+\sqrt{5}}{2}, w_2 = \frac{1-\sqrt{5}}{2}$$

For any A_1 substituting $A_1 w_1^n$ in $u_n - u_{n-1} - u_{n-2}$ yields zero. For any A_2 substituting $A_2 w_2^n$ for u_n in $u_n - u_{n-1} - u_{n-2}$ yields zero. And this suggests a general solution as,

$$u_n = A_1 w_1^n + A_2 w_2^n$$

Check this by substituting into $u_n - u_{n-1} - u_{n-2}$ thus:

$$(A_1 w_1^n + A_2 w_2^n) - (A_1 w_1^{n-1} + A_2 w_2^{n-1}) - (A_1 w_1^{n-2} + A_2 w_2^{n-2})$$

This, arranged, is

$$A_1 w_1^{n-2}(w_1^2 - w_1 - 1) + A_2 w_2^{n-2}(w_2^2 - w_2 - 1)$$

And this is zero because both expressions in brackets are zero. On substituting the values of w_1 and w_2 the general solution is,

$$u_n = A_1(\frac{1+\sqrt{5}}{2})^n + A_2(\frac{1-\sqrt{5}}{2})^n$$

but, by noting initial conditions, values for A_1 and A_2 can be obtained. And note that,

$$u_0 = 1 \Rightarrow A_1 + A_2 = 1 \& A_2 = 1 - A_1$$

Likewise,

$$u_1 = 1 \Rightarrow \frac{A_1(1+\sqrt{5}) + (1-A_1)(1-\sqrt{5})}{2} = 1$$

so

$$A_1 = \frac{1+\sqrt{5}}{2\sqrt{5}}$$

And $A_2 = 1 - A_1 = 1 - \frac{1+\sqrt{5}}{2\sqrt{5}} = \frac{1-\sqrt{5}}{2\sqrt{5}}$ In consequence,

$$u_n = \frac{1+\sqrt{5}}{2\sqrt{5}}(\frac{1+\sqrt{5}}{2})^n - \frac{1-\sqrt{5}}{2\sqrt{5}}(\frac{1-\sqrt{5}}{2})^n$$

Thus

$$u_n = \frac{1}{\sqrt{5}}\left[(\frac{1+\sqrt{5}}{2})^{n+1} - (\frac{1-\sqrt{5}}{2})^{n+1}\right]$$

as the final solution.

4 Second Order Homogeneous Linear Difference Equation — II

To solve

$$u_n = p \cdot u_{n+1} + q \cdot u_{n-1}, with, u_0 = 1, u_l = 1, and, p + q = 1$$

Transfer all the terms to the left-hand side:

$$p \cdot u_{n+1} - u_n + q \cdot u_{n-1} = 0$$

Guess:

$$u_n = A \cdot w^n$$

so

$$pAw^{n+1} - Aw^n + qAw^{n-1} = 0$$
$$\Rightarrow pw^2 - 2 + q = 0$$
$$\Rightarrow pw^2 - (p+q)w + q = 0$$
$$\Rightarrow (w-1)(pw-q) = 0$$

Then two roots are

$$w_1 = 1, and, w_2 = \frac{q}{p}$$

This suggests a general solution:

$$u_n = A_1(1)^n + A_2(\frac{q}{p})^n, provided, p \neq q$$

Check by substituting into $q \cdot u_{n+1} - u_n + q \cdot u_{n-1}$, thus

$$\left[pA_1(1)^{n+1} + pA_2(\frac{q}{p})^{n+1}\right] - \left[A_1(1)^n + A_2(\frac{q}{p})^n\right] + \left[qA_1(1)^{n-1} + qA_2(\frac{q}{p})^{n-1}\right]$$

This, rearranged, is:

$$A_1\left[p - 1 + q\right] + A_2(\frac{q}{p})^{n-1}\left[p(\frac{q}{p})^2 - \frac{q}{p} = q\right]$$

which, given that $p + q = 1$, is

$$A_2(\frac{q}{p})^{n-1}\left[\frac{q^2}{p} - \frac{q}{p} + q\right] = A_2(\frac{q}{p})^{n-1}\left[\frac{q}{p}(q-1) + q\right] = A_2(\frac{q}{p})^{n-1}\left[\frac{q}{p}(-p) + q\right] = 0$$

The next step is to determine values for A_1 and A_2 for general solution. And the general solution was determined to be,

$$u_n = A_1(1)^n + A_2(\frac{q}{p})^n, with, p \neq q$$

Note

$$u_0 = 0 \Rightarrow A_1 + A_2 = 0$$

Likewise

$$u_l = 1 \Rightarrow A_1 + A_2(\frac{q}{p})^l = 1$$

so

$$-A_2 + A_2(\frac{q}{p})^l = 1 \Rightarrow A_2 = \frac{1}{(\frac{q}{p})^l - 1}$$

and

$$A_1 = -A_2 = \frac{-1}{(\frac{q}{p})^l - 1}$$

In consequence:

$$u_n = \frac{-1}{(\frac{q}{p}) - 1} + \frac{(\frac{q}{p})^n}{(\frac{q}{p})^l - 1}$$

giving

$$u_n = \frac{(\frac{q}{p})^n - 1}{(\frac{q}{p})^l - 1}$$

as the final solution.

Remark 4.1. *First,* $u_0 = 0$*, and* $u_1 = 1$ *as required. Second, suppose* $0 \ll n \ll l$ *(e.g.: l=57, and n=41)...*

- *If* $\frac{q}{p} < 1$, $\left[(\frac{q}{p})^i \rightarrow 0 for large i\right]$*, the solution* $u_n \rightarrow \frac{0-1}{0-1} \rightarrow 1$
- $\frac{q}{p} > 1$*, the solution* $u_n \rightarrow \frac{(\frac{q}{p})^n\left[1-(\frac{p}{q})^n\right]}{(\frac{q}{p})^l\left[1-(\frac{p}{q})^l\right]} \rightarrow \frac{1}{(\frac{q}{p})^{l-n}}\left[\frac{1-0}{1-0}\right] \rightarrow 0$

In simple terms, provided n is well clear of the extremes 0 and l, u_n will tend to 1 or to 0 depending on whether $q < p$ or $q > p$. (It has been assumed that $p \neq q$)

Remark 4.2. (what about the case $p = q$ as for an even coin) Recall that $w_1 = 1$, and $w_2 = \frac{q}{p}$ so the case $p = q$ implies twin roots, $w_1 = w_2 = 1$. The general solution $u_n = A_1 w_1^n + A_2 w_2^n$ would be $u_n = A_1 + A_2$ which is silly. In such case, try a different guess:

$$u_n = (A_1 + A_2 n) w^n$$

where w is the twin root. In the present case, try

$$u_n = (A_1 + A_2 n)(1)^n$$

as the general solution. Check by substituting into $p \cdot u_{n+1} - u_n - q \cdot u_{n-1}$, thus:

$$p\left[A_1 + A_2(n+1)\right] - \left[A_1 + A_2 n\right] + q\left[A_1 + A_2(n-1)\right]$$

This, arranged, is:

$$A_1\left[p - 1 + q\right] + A_2\left[pn + p - n + qn - q\right]$$

which, remembering that $p + q = 1$, is zero

The next step is to determine values for A_1 and A_2 in the general solution whose revised form is

$$u_n = (A_1 + A_2 n)(1)^n$$

Note

$$u_0 = 0 \Rightarrow A_1 = 0$$

Likewise

$$u_l = 1 \Rightarrow 0 + A_2 l = 1, if, A_2 = \frac{1}{l}$$

In consequence

$$u_n = 0 + \frac{1}{l} n$$

giving

$$u_n = \frac{n}{l}$$

as the final solution when the special case $p = q$ applies.

5 Second Order Inhomogeneous Linear Difference Equation

To solve

$$v_n = 1 + p v_{n+1} + q v_{n-1}$$

given that $v_0 = v_l = 0$, and $p + q = 1$, and transfer all the terms except the 1 to the left-hand side:

$$p v_{n+1} - v_n + q v_{n-1} = -1$$

If the right-hand side were zero, this would be identical to the homogeneous equation just discussed. The new equation is solved in two steps. First, deem the right-hand side to be zero and solve as for the homogeneous case:

$$v_n = A_1(1)^n + A_2(\frac{q}{p})^n$$

provided that $q \neq p$. Then, augment this solution by some $f(n)$ which has to be given further thought:

$$v_n = A_1(1)^n + A_2(\frac{q}{p})^n + f(n)$$

This augmented v_n has to be such that when substituted into $pv_{n+1} - v_n + qv_{n-1}$ the result is -1.

From previous experience with u_n, it is known that substituting $A_1(1)^n + A_2(\frac{q}{p})^n$ gives a result of zero. In consequence, the property required of $f(n)$ is that on substituting it into $pv_{n+1} - v_n + qv_{n-1}$ the result must be -1

It will always be reasonable to express $f(n)$ as the quadratic $a + bn + cn^2$ with only one of the constants a, b and c non-zero. In the present case tray $f(n) = kn$ and therefore require:

$$pk(n+1) - kn + qk(n-1) = -1$$

so

$$pkn + pk - kn = qkn - qk = -1$$

Hence $(p - q)k = -1$ so $k = \frac{1}{q-p}$, giving:

$$v_n = A_1 + A_2(\frac{q}{p})^n + \frac{n}{q-p}$$

as the general solution appropriate to the inhomogeneous difference equation. It is left as an exercise for the reader to determine values for $A(1)$ and $A(2)$ appropriate for the initial conditions given.

Remark 5.1. (What about the case $p = q$?) When $p = q$ the equation $pv_{n+1} - v_n + qv_{n-1} = -1$ can be solved in two steps as before. First, deem the right-hand side to be zero and solve as for the homogeneous case:

$$v_n = (A_1 + A_2 n)(1)^n$$

Then, augment this solution by some $f(n)$ which has to be given further thought:

$$v_n = (A_1 + A_2 n)(1)^n + f(n)$$

As before, this augmented v_n has to be such that when substituted into $pv_{n+1} - v_n + qv_{n-1}$ the results is -1 but remember that $p = q$ this time. Again, from previous experience with un, it is known that substituting $(A_1 + A_2 n)(1)^n$ gives a result of zero. Once more, the property required of $f(n)$ is that on substituting it into $pv_{n+1} - v_n + qv_{n-1}$ the result must be -1.

Since $p = q$, it is no use this time employing the previous approach which was to try $f(n) = kn$ and derive $k = \frac{1}{q-p}$. This is not a helpful value for k!

The appropriate approach now is to try $f(n) = kn^2$ and require

$$pk(n+1)^2 - kn^2 + qk(n-1)^2 = -1$$

so

$$pkn^2 + 2pkn + pk - kn^2 + qkn^2 - 2qkn + qk = -1$$

Hence $(p+q)k = -1$ so $k = -1$, giving

$$v_n = A_1 + A_2 n - n^2$$

as the general solution appropriate to the inhomogeneous difference equation when $p = q$. Note that $A_1 + A_2 n$ is the solution to the homogeneous equation when $p = q$ and $-n^2$ is the required augmentation. Given the initial conditions $v_0 = v_l = 0$, it is easy to determine that $A_1 = 0$, and $A_2 = l$ giving:

$$v_n = n(l - n)$$

as the final solution when the special case $p = q$ applies.

6 Lag Operator

Definition 6.1. (Lag Operator) Denote L, and $L \cdot y_t = y_{t-1}$

Remark 6.2.

- *The lag operator could be raised to powers, e.g. $L^2 y_t = y_{t-2}$. We could also form polynomials of it*

$$a(L) = a_0 + a_1 L + a_2 L^2 + ... + a_p L^p$$

$$a(L) y_t = a_0 y_t + a_1 y_{t-1} + a_2 y_{t-2} + ... + a_p y_{t-p}$$

- *Lag polynomials could be multiplied. Multiplication is commutative, $a(L)b(L) = b(L)a(L)$*
- *Some lag polynomials could be inverted. We define $(1 - \rho L)^{-1}$ by the following equality*

$$(1 \rho L)(1 - \rho L)^{-1} \equiv 1$$

Theorem 6.3. *If $|\rho| < 1$, then*

$$(1 - \rho L)^{-1} = \sum_{i=0}^{\infty} \rho^i L^i$$

Proof:

$$(1 - \rho L) \sum_{i=0}^{\infty} \rho^i L^i = \sum_{i=0}^{\infty} \rho^i L^i - \sum_{i=1}^{\infty} \rho^i L^i = \rho^0 L^0 = 1$$

For higher order polynomials, we can invert them by factoring, using the formula for $(1 - \rho L)^{-1}$ (assuming that the roots are outside the unit circle), and then rearranging, for example:

$$1 - a_1 L - a_2 L^2 = (1 - \lambda_1 L)(1 - \lambda_2 L), with, |\lambda_i| < 1$$

$$(1 - a_1 L - a_2 L^2)^{-1} = (1 - \lambda_1 L)^{-1} (1 - \lambda_2 L)^{-1}$$

$$\Rightarrow (1 - a_1 L - a_2 L^2)^{-1} = (\sum_{i=0}^{\infty} \lambda_1^i L^i)(\sum_{i=0}^{\infty} \lambda_2^i L^i) = \sum_{j=0}^{\infty} L^j (\sum_{k=0}^{j} \lambda_1^k \lambda_2^{j-k})$$

Another (perhaps more easy) way to approach the same problem is do a partial fraction decomposition

$$\frac{1}{(1-\lambda_1 x)(1-\lambda_2 x)} = \frac{a}{1-\lambda_1 x} + \frac{b}{1-\lambda_2 x}$$

$$\Rightarrow a = \frac{\lambda_1}{\lambda_1 + \lambda_2}, b = \frac{\lambda_2}{\lambda_1 + \lambda_2}$$

$$a^{-1}(L) = a \sum_{i=0}^{\infty} \lambda_1^i L^i + b \sum_{i=0}^{\infty} \lambda_2^i L^i$$

This trick only works when the λ_i are unique. The formula is slightly different otherwise.

7 Application I: Simple Process in Time Series

7.1 Autoregressive (AR)

$$AR(1) : y_t = \rho y_{t-1} + e_t, |\rho| < 1$$

$$(1 - \rho L) y_t = e_t$$

$$AR(p) : a(L) y_t = e_t$$

Where $a(L)$ is of order p

7.2 Moving Average (MA)

$$MA(1) : y_t = e_t + \theta e_{t-1}$$

$$y_t = (1 + \theta L) e_t$$

$$MA(q) : y_t = b(L) e_t$$

where $b(L)$ is of order q

7.3 ARMA

$$ARMA(p, q) : a(L) y_t = b(L) e_t$$

where $a(L)$ is order p and $b(L)$ is order q, and $a(L)$ and $b(L)$ are relatively prime

An ARMA representation is not unique, For example, an $AR(1)$ (with $|\rho| < 1$) is equal to an $MA(\infty)$, as we saw above. We can see it from the formula for inversion:

$$y_t = \sum_{j=0}^{\infty} \rho^j e_{t-j}$$

Aside, you can also get this formula by repeatedly using definition of $AR(1)$, as

$$y_t = \rho y_t + e_t = \rho(\rho y_{t-2} + e_{t-1}) + e_t = ... = \sum_{j=0}^{k-1} \rho^j e_{t-j} + \rho^k y_{t-k}$$

and noticing that $\rho^k \overset{\mathcal{L}^2}{\rightarrow} 0$ as $k \rightarrow \infty$. In fact, this is more generally true. Any $AR(p)$ with roots outside the unit circle has an MA representation. These processes are called stationary (Because there is weakly stationary version of them).

Any MA process with roots outside unit circle could also be written as $AR(\infty)$, such process called invertible. If $y_t = b(L)e_t$ is an invertible MA process, then $e_t = b(L)^{-1}y_t$. That is, the "errors" are laying in a space of observations and could be recovered from y's (another name for this: errors are fundamental).

8 Application II: From *Stochastic Differential Equation* to *Autoregressive Time Series Model*

Remind ourselves of ordinary differential equation that has been studied,

$$\frac{dY(t)}{dt} = -\kappa Y + b(t) = p(t)Y + b(t), with, Y(0) = 0$$

where $p(t) = -\kappa$ but it could be more general. Remind ourselves of the *Integrating Factor*, which is $v(t)$ defined as,

$$v(t) = exp(-\int_0^t p(r)dr) = exp(\kappa t)$$

By definition, $dv(t)/dt = -v(t)p(t) = \kappa v(t)$. Consider

$$\frac{d(v(t)Y(t))}{dt} = v(t)\frac{dY(t)}{dt} + Y(t)\frac{dv(t)}{dt} + Y(t)\frac{dv(t)}{dt} = v(t)b(t)$$

The solution is

$$v(t)Y(t) - v(0)Y(0) = \int_0^t v(r)b(r)dr = \int_0^t exp(\kappa r)b(r)dr$$

$$Y(t) = Y_0 exp(-\kappa t) + \int_0^t exp(-\kappa(t-r))b(r)dr$$

with $Y(0) = 0$, then

$$Y(t) = \int_0^t exp(-\kappa(t-r))b(r)dr$$

Thus the following stochastic differential equation could be solved using the similar manner, which is

$$dY(t) = -\kappa Y dt + \sigma dB(t), Y(0) = 0$$

And the general solution could be derived as,

$$Y(t) = \int_0^t \sigma exp(-\kappa(t-r))dB(r)$$

and we could discretize the process as,

$$Y(t+h) = e^{-\kappa h}Y(t) + \int_0^h \sigma exp(-\kappa(h-s))dB(t+s)$$

This is an *AR(1)* model. Since

$$\int_0^h \sigma e^{-\kappa(h-s)}dB(t+s)\ N(0, \int_0^h \sigma^2 e^{-2\kappa(h-s)}ds) = N(0, \frac{\sigma^2(1-exp(-2\kappa h))}{2\kappa})$$

Then the transition density is

$$Y(t+h)|Y(t)\ N(e^{-\kappa h}Y(t), \frac{\sigma^2(1-exp(-2\kappa h))}{2\kappa})$$

Besides deriving the transition probability from the AR process, we could also discretize the stochastic differential equation into autoregressive regression model to finish the estimation, and the procedure is conducted as,

$$dY(t) = -\kappa Y(t)dt + \sigma dB(t)$$

$$\Rightarrow Y(t+1) - Y(t) = -\kappa Y(t)\left[(t+1) - t\right] + \sigma N(0,1)$$
$$\Rightarrow Y(t+1) = (1-\kappa)Y(t) + \sigma N(0,1)$$
$$\Rightarrow Y(t+1) = \rho Y(t) + \epsilon(t)$$

Where $\rho = 1-\kappa$, and $\epsilon(t) = \sigma N(0,1)$ And using the techniques of OLS, we could derive $\widehat{\rho}$, and if the interval is going to zero, then $\widehat{\rho} \xrightarrow{p} 1-\kappa$ according to law of large number.

Chapter V: Partial Differential Equation & Stochastic Differential Equation

Liu Yanbo

May 24, 2018

Abstract

Some of the stochastic differential equation could not easily be solved using simple method for ordinary differential equations, so more advanced methods as partial differential equation are in need. This chapter is based on the following materials. This chapter is based on the following teaching materials.

- Strauss, Walter A. Partial differential equations. Vol. 92. New York: Wiley, 1992.
- Øksendal, Bernt. "Stochastic differential equations." Stochastic differential equations. Springer Berlin Heidelberg, 2003.
- Lecture Notes of "Introduction to PDE", National Chiao Tung University, http://ocw.nctu.edu.tw/
- Lecture Notes of "Introduction to Financial Mathematics II", National Chiao Tung University, http://ocw.nctu.edu.tw/

1 Mathematical Preliminary I: Laplace Transform

- Transforms: An operation that transform a function into another function
  - Differentiation transform
  $$\frac{d}{dx}(x^2) = 2x$$
  - Integration transform:
  $$\int x^2 dx = \frac{1}{3}x^3 + c$$
- Now consider a defined integral $\int_0^\infty k(s,t)f(t)dt$ that transforms $f(t)$ into a function of variables
  - The integral is said to be convergent if the limit exists:
  $$\int_0^\infty k(s,t)f(t)dt = lim_{b\to\infty}\int_0^b k(s,t)f(t)dt$$

- The integral is said to be divergent if the limit does not exist

- Laplace transform: one kind of integration transform
  - Definition: $L[f(t)] = \int_0^\infty e^{-st} f(t)dt$, where $f(t)$ is a function defined for $t \geq 0$, and s,t are two independent variables.
  - When the above integral converges, the result is a function of s:

$$L[f(t)] = \int_0^\infty e^{-st} f(t)dt$$

Remark 1.1. *$L[1] \equiv \int_0^\infty e^{-st}dt = lim_{b\to\infty} e^{-st}dt = lim_{b\to\infty} \frac{-e^{-st}}{s}|_0^b = lim_{b\to\infty} \frac{-b^{-sb}}{+} 1s = \frac{1}{2}$ for $s > 0$; similarly we could derive more results, as*
$L[1] = \frac{1}{s}, for, s > 0$
$L[t^n] = \frac{n!}{s^{n+1}}, \forall n > -1, s > 0$
$L[e^{at}] = \frac{1}{s-1}, (s-a) > 0$
$L[sin(kt)] = \frac{k}{s^2+k^2}, s > 0$
$L[cos(kt)] = \frac{s}{s^2+k^2}, for, s > 0$

Remark 1.2.
- *Linear property $L[\alpha f(t) + \beta g(t)] = \alpha L[f(t)] + \beta L[g(t)] = \alpha F(s) + \beta G(s)$*
- *Existence of $L[f(t)]$,*
  - *Sufficient condition 1: for the existence of $L[f(t)]$, the graph of $f(t)$ cannot grow faster than the graph of est as t increase.*
  - *Sufficient condition 2: $f(t)$ can be a stepwise continuous function on the interval $[0, \infty)$*

Definition 1.3. *If $F(s)$ represents the L-transform of $f(t)$, i.e. $L[f(t)] = F(s)$, we then say $f(t)$ is the inverse L-transform of $F(s)$, and could be expressed as follows:*

$$L^{-1}[F(s)] = \frac{1}{2\pi i} \int_{r-i\infty}^{r+i\infty} e^{st} F(s)ds = f(t), \forall t \geq 0, r \in R$$

Remark 1.4. $L^{-1}\left[\frac{1}{2}\right] = 1$
$L^{-1}\left[\frac{n!}{s^{n+1}}\right] = t^n$
$L^{-1}\left[\frac{1}{s-1}\right] = e^{at}$
$L^{-1}\left[\frac{k}{s^2+k^2}\right] = sin(kt)$
$L^{-1}\left[\frac{s}{s^2+k^2}\right] = cos(kt)$
$L^{-1}\left[\frac{k}{s^2-k^2}\right] = sinh(kt)$
$L^{-1}\left[\frac{s}{s^2-k^2}\right] = cosh(kt)$
And the linearity also holds here.

2 Mathematical Preliminary II: Fourier Transformation

Definition 2.1. *The Fourier expansion of f defined on the interval $(-P, P)$ is given by*

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty}\left[a_n cos\frac{n\pi}{p}x + b_n sin\frac{n\pi}{p}x\right]$$

where $a_0 = \frac{1}{p}\int_{-p}^{p} f(x)dx$
$a_n = \frac{1}{p}\int_{-p}^{p} f(x)cos(\frac{n\pi}{p}x)dx$
$b_n = \frac{1}{p}\int_{-p}^{p} sin(\frac{n\pi}{p}x)dx \rightarrow$ *Fourier coefficient of f.*

Example 2.2. *Expand* $f(x) = 0, -\pi < x < 0$ *and* $f(x) = 2 - x, 0 \leq x < \pi$ *in a Fourier series*
Hence $(-p, p) = (-\pi, \pi)$
$a_0 = \frac{1}{p}\int_{-p}^{p} f(x)dx = \frac{\pi}{2}$
$a_n = \frac{1}{p}\int_{-p}^{p} f(x)cos(\frac{n\pi}{p}x)dx = \frac{1-(-1)^n}{n^2\pi}$
$b_n = \frac{1}{p}\int_{-p}^{p} sin(\frac{n\pi}{p}x)dx = \frac{1}{n}$
$\Rightarrow f(x) = \frac{\pi}{4} + \sum_{n=1}^{\infty}\left[\frac{1-(-1)^n}{n^2\pi}cos(nx) + \frac{1}{n}sin(nx)\right]$

Remark 2.3. (Cosine and Sine series)

- If f has $f(-x) = f(x)$ property on $(-P, P)$, then its Fourier expansion is
  $a_0 = \frac{1}{p}\int_{-p}^{p} f(x)dx = \frac{2}{p}\int_{0}^{p} f(x)dx$
  $a_n = \frac{1}{p}\int_{-p}^{p} f(x)cos(\frac{n\pi}{p}x)dx = x\int_{0}^{p} cos(\frac{n\pi}{p}x)dx$
  $b_n = \frac{1}{p}\int_{-p}^{p} sin(\frac{n\pi}{p}x)dx = 0 \Rightarrow f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n cos(\frac{n\pi}{p}x)$
- If f has $f(-x) = -f(x)$ property on $(-P, P)$, then its Fourier expansion is

  $$f(x) = \sum_{n=1}^{\infty} b_n sin(\frac{n\pi}{p}x) \rightarrow Fouriersineseries$$

  where $b_n = \frac{2}{p}\int_{0}^{p} f(x)sin(\frac{n\pi}{p}x)dx$

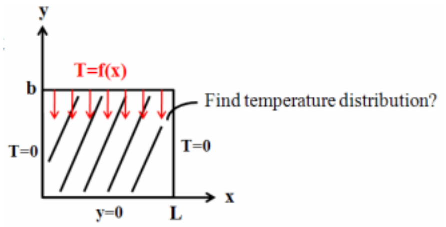
Example 2.4. *Expand* $f(x) = -1, -\pi < x < 0$ *and* $f(x) = 1, 0 < x < \pi$ *in a Fourier series. Note that the interval* $(-p, p) = (-\pi, \pi)$
Since $f(x)$ *is odd at* $(-\pi, \pi)$, $f(x)$ *could be simply expanded in sine series*

$$\Rightarrow f(x) = \sum_{n=1}^{\infty} b_n sin(\frac{n\pi}{p}x)$$

where $b_n = \frac{2}{p}\int_{0}^{p} f(x)sin(\frac{n\pi}{p})dx = \frac{2}{\pi}\int_{0}^{\pi} f(x)sin(nx)dx = \frac{2}{\pi}\cdot\frac{1-(-1)^n}{n}$

Remark 2.5. *In general,* $f(x)$ *defined on* $(a, a + 2p), a \in \mathbb{R}$ *could also be expanded in a Fourier series. Suppose that* $f(x)$ *is a function defined on* $(0, 2p)$ *and Fourier series is given by*

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty}\left[a_n cos(\frac{n\pi}{p}x) + b_n sin(\frac{n\pi}{p}x)\right]$$



where $a_o = \frac{1}{p}\int_0^{2p} f(x)dx$
$a_n = \frac{1}{p}\int_0^{2p} f(x)cos(\frac{n\pi}{p}x)dx$
$b_n = \frac{1}{p}\int_0^{2p} f(x)sin(\frac{n\pi}{p}x)dx$

3 Elementary Partial Differential Equation

Here we only focus on how to solve the simple partial differential equations, and three main solving techniques are provided,
Separation of Variable
Laplace Transformation
Combination of Variable

3.1 Separation of Variable Method

The method of Separation of Variables cannot always be used and even when it can be used it will not always be possible to get much past the first step in the method. However, it can be used to easily solve the $1 - D$ heat equation with no sources, the $1 - D$ wave equation, and the $2 - D$ version of Laplace's Equation, $\bigtriangledown^2 u = 0$.

In order to use the method of separation of variables we must be working with a linear homogenous partial differential equations with linear homogeneous boundary conditions. At this point we're not going to worry about the initial condition(s) because the solution that we initially get will rarely satisfy the initial condition(s). As we'll see however there are ways to generate a solution that will satisfy initial condition(s) provided they meet some fairly simple requirements.

The method of separation of variables relies upon the assumption that a function of the form,

$$u(x,t) = \psi(x)G(t)$$

will be a solution to a linear homogeneous partial differential equation in x and t. This is called a product solution and provided the boundary conditions are also linear and homogeneous this will also satisfy the boundary conditions. However, as noted above this will only rarely satisfy the initial condition, but that is something for us to worry about in the next section.

Example 3.1. *Give* $T(x,y)$, *a function of x&y: steady-state, solid and no heat generation* $\Rightarrow \bigtriangledown^2 T = 0 \Rightarrow \frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial^2} = 0$ *Boundary Conditions are,*
$x = 0, T = 0$
$y = 0, T = 0$

$x = L, T = 0$
$y = b, T = f(x)$

Let $T(x,y) = \overline{X}(x) \cdot \overline{Y}(y)$ *be put back into the original equation* $\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial^2} = 0 \Rightarrow \overline{Y}(y)\frac{d^2\overline{X}}{dx^2} + \overline{X}(x)\frac{d^2\overline{Y}}{dy^2} = 0$

$$\Rightarrow \frac{1}{\overline{X}}\frac{d^2\overline{X}}{dx^2} + \frac{1}{\overline{Y}}\frac{d^2\overline{Y}}{dy^2} = 0$$

$$\Rightarrow -\frac{1}{\overline{X}}\frac{d^2\overline{X}}{dx^2} = \frac{1}{\overline{Y}}\frac{d^2\overline{Y}}{dy^2} = \lambda^2, \lambda : Eigenvalueconstant$$

$$\begin{cases} \frac{d^2\overline{X}}{dx^2} + \lambda^2\overline{X} = 0 & x = 0, T = 0 \to T(0,y) = \overline{X}(0)\cdot\overline{Y}(y) = 0 \\ & x = L, T = 0 \to T(L,y) = \overline{X}(L)\cdot\overline{Y}(y) = 0 \\ \frac{d^2\overline{Y}}{dy^2} + \lambda^2\overline{Y} = 0 & y = 0, T = 0 \to \overline{Y}(y) = 0 \\ & y = b, T = f(x) \end{cases}$$

- *From* $\frac{d^2\overline{X}}{dx^2} + \lambda^2\overline{X} = 0$ *(2nd order homogeneous ordinary differential equation)*
  Let $\overline{X} = e^r$, $r^2 + \lambda^2 = 0$,
  $r^2 = i\lambda$ *or* $r^2 = -i\lambda$,
  $\overline{X} = Acos\lambda x + Bsin\lambda x$
  Put $BC1$ *into the system, and we have*
  $A\overline{Y}(y) = 0 \Rightarrow A = 0$
  Put $BC2$ *into the system, and we have*
  $Bsin\lambda L = 0$, *and* $B \neq 0$, *thus* $sin\lambda L = 0 \Rightarrow \lambda = \frac{n\pi}{L}, n = 1,2,3,...$*(eigen value)*
  so $\overline{X} = Bsin(\frac{n\pi}{L}x)$*(eigen function)*

- *From* $\frac{d^2\overline{Y}}{dy^2} + \lambda^2\overline{Y} = 0, r^2 - \lambda^2 = 0 \Rightarrow r^2 = \pm\lambda, \overline{Y} = Ce^{\lambda y} + De^{-\lambda y}$
  and since $cosh(\lambda y) = \frac{e^{\lambda y}+e^{-\lambda y}}{2}$, $sinh(\lambda y) = \frac{e^{\lambda y}-e^{-\lambda y}}{2}$

$$\Rightarrow \overline{Y} = C\cdot cosh(\lambda y) + D\cdot sinh(\lambda y)$$
$$\Rightarrow \overline{Y} = Ccosh(\frac{n\pi}{L}y) + Dsinh(\frac{n\pi}{L}y)$$

  Put $BC3$ *into the system, and we have*
  $0 = C\cdot cosh0 + D\cdot sinho = C \Rightarrow \overline{Y} = S\cdot sinh(\frac{n\pi}{L}y)$

- $T(x,y) = \overline{X}(x)\cdot\overline{Y}(y) = \sum_{n=1}^{\infty} Bsin(\frac{n\pi}{L}x)\cdot Dsinh(\frac{n\pi}{L}y) = \sum_{n=1}^{\infty} E_n sin(\frac{n\pi}{L}x)\cdot Dsinh(\frac{n\pi}{L}y)$

- *Put* $BC4$ *into* $f(x) = \sum_{n=1}^{\infty} E_n sinh(\frac{n\pi b}{L})\cdot sinh(\frac{n\pi x}{L})$
  By the definition of Fourier sin-series expansion, we have

$$E_n sinh(\frac{n\pi}{L}b) = \frac{L}{2}\int_0^L f(x)sin(\frac{n\pi x}{L})dx$$
$$\Rightarrow E_n = \frac{\frac{L}{2}\int_0^L f(x)sin(\frac{n\pi x}{L})dx}{sinh(\frac{n\pi b}{L})}$$

- $T(x,y) = \sum_{n=1}^{\infty} E_n sin(\frac{n\pi x}{L})\cdot sinh(\frac{n\pi y}{L})$, *where* $E_n = \frac{\frac{L}{2}\int_0^L f(x)sin(\frac{n\pi x}{L})dx}{sinh(\frac{n\pi b}{L})}$

3.2 Laplace Transformation Method

Assume $T(x,t), x \leq x \leq b, t \geq 0, L\left[T(x,t)\right] = T(x,s)$
1.$L\left[\frac{\partial T}{\partial t}\right] = s\overline{T}(x,s) - \overline{T}(x,0)$
2.$L\left[\frac{\partial^2 T}{\partial t^2}\right] = s^2\overline{T}(x,s) - s\overline{T}(x,0)$, where $T_t = \frac{\partial T}{\partial t}$
3.$L\left[\frac{\partial T}{\partial x}\right] = \frac{d}{dx}\overline{T}(x,s)$
4.$L\left[\frac{\partial^2 T}{\partial x^2}\right] = \frac{d^2}{dx}\overline{T}(x,s)$

Example 3.2. $\frac{\partial T}{\partial x} + x\frac{\partial T}{\partial t} = 0, T(x,0) = 0, T(0,t) = 4t$

Take Laplace on each part of PDE, which is

$$\frac{dT}{dx} + x\left[s\overline{T} - \overline{T}(x,0)\right]$$

And put $T(x,0) = 0$ *into it then get* $\frac{dT}{dx} + xs\overline{T} = 0(1^{st} orderlinearODE)$
$\overline{T}(x,s) = c \cdot exp(-\frac{1}{2}sx^2), \overline{T}(0,s) = L\left[T(0,t)\right] = L\left[4t\right] - \frac{4}{s^2} \Rightarrow c = -\frac{4}{s^2} \Rightarrow$ $\overline{T}(x,s) = -\frac{4}{s^2} \cdot exp(-\frac{1}{2}sx^2)$ *Thus the solution for PDE is as*

$$\overline{T}(x,t) = L^{-1}\left[T(x,s)\right] = L^{-1}\left[-\frac{4}{s^2} \cdot exp(-\frac{1}{2}sx^2)\right] = 4(t - \frac{1}{2}x^2)u(t - \frac{1}{2}x^2)$$

By 2^{nd}*−transition theorem*

Remark 3.3. *Other kinds of PDE could be solved by Laplace transformation*

- *Heat equation* $k\frac{\partial u^2}{\partial x^2} = \frac{\partial u}{\partial t}$
- *Laplace equation* $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$
-

3.3 Combination of Variable Method

Remark 3.4. *Combination of Variable method is useful for* homogeneous ODE, *and* non-homogenous Boundary Condition

Example 3.5. *Given* $Q(x,t)$ *a function of* $x\&t$,

$$\frac{\partial Q}{\partial t} = \alpha\frac{\partial^2 Q}{\partial x^2}$$

And the Boundary Condition is given as,

$$t = 0, Q = 0; x = 0, Q = 1; x = \infty, Q = 0$$

By combination of variable method,
Let $\eta = ax^b t^c \Rightarrow Q(x,t) = Q(\eta)$
Then $\frac{\partial Q}{\partial t} = \frac{dQ}{d\eta} \cdot \frac{\partial \eta}{\partial t} = \frac{c\eta}{t}Q'$
$\frac{c\eta}{}t = cax^b t^{c-1} = c\frac{\eta}{t}$
$\frac{\partial Q}{\partial x} = \frac{dQ}{d\eta} \cdot \frac{\partial \eta}{\partial x} = \frac{b\eta}{x}Q'$

$\frac{\partial^2 Q}{\partial x^2} = \frac{\partial}{\partial x}(\frac{\partial Q}{\partial x}) = \frac{-b\eta Q'}{x^2} + \frac{b^2\eta}{x^2}(\eta Q')'$
And put them back into the original ODE $\frac{\partial Q}{\partial t} = \alpha\frac{\partial^2 Q}{\partial x^2}$

$$\Rightarrow \frac{c\eta}{t}Q = \alpha\left[\frac{-b\eta Q'}{x^2} + \frac{b^2\eta}{e^2}(\eta Q')'\right]$$

Multiply both sides by $\frac{x^2}{\eta}$

$$\frac{cx^2}{t}Q' = \alpha\left[-bQ' + b^2(\eta Q')'\right] = \alpha\left[-bQ' + b^2Q' + b^2\eta Q''\right]$$

Because of the relationship of $\frac{x^2}{t}$, $b : c = 2 : (-1)$, *and* $b^2 - b = 0 \Rightarrow b = 1$

$$\Rightarrow -\frac{1}{2}(\frac{\eta}{a})^2 Q'\alpha\left[-Q' + Q' + \eta Q''\right]$$

$$\Rightarrow Q'' + \frac{\eta}{2a^2\alpha}Q' = 0$$

Replace $Q' = u$, *and* $Q'' = \frac{du}{d\eta}$

$$\Rightarrow \frac{du}{d\eta} + \frac{\eta}{2a^2\alpha}u = 0$$

which is 1^{st}*-order ODE, and take integration on both sides of the equation, which is as,*

$$ln(u) + \frac{\eta^2}{4a^2\alpha} = ln(c_1)$$

And let $\frac{1}{4a^2\alpha} = 1 \Rightarrow a = \frac{1}{\sqrt{4\alpha}}$
Thus

$$\Rightarrow \eta = ax^bt^c = \frac{x}{\sqrt{4\alpha t}}$$

$$\Rightarrow ln(u) + \eta^2 = ln(c_1)$$

$$\Rightarrow u = c_1e^{-\eta^2} = \frac{dQ}{d\eta}$$

$$\Rightarrow Q = c_1\int_0^\eta d^{-\eta^2}d\eta + c_2$$

Boundary conditions are $\eta = 0, Q = 1$, *and* $\eta = \infty, Q = 0$ *Put* $BC1$ *into* $Q = c_1\int_0^\eta d^{-\eta^2}d\eta + c_2 \Rightarrow c_2 = 1$ *And Put* $BC2$ *into* $Q = c_1\int_0^\eta d^{-\eta^2}d\eta + c_2 \Rightarrow c_1 = \frac{1}{\int_0^\eta e^{-\eta^2}d\eta}$
Finally, the solution is

$$Q = 1 - \frac{\int_0^\eta e^{-\eta^2}d\eta}{\int_0^\infty e^{-\eta^2}d\eta}$$

4 Feynman-Kac formula

Theorem 4.1. (Feynman-Kac formula) Consider the stochastic differential equation

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

Let f be a Borel-measurable function. Fixe $T > 0$ and let $t \in [0, T]$ be given. Define the function,

$$g(t, x) = \mathbb{E}^{tx}\left[f(X_T)\right] = \mathbb{E}\left[f(X_T)|X_T = x\right]$$

Assume that $g(t, x) < \infty$ for all (t, x). Then $g(t, x)$ satisfies the partial differential equation

$$g_t(t, x) + \beta(t, x)g_x(t, x) + \frac{1}{2}\gamma^2(t, x)g_{xx}(t, x) = 0$$

with the terminal condition $g(T, x) = f(x)$ for all x.

Remark 4.2. $(g(t, X_t))_{0 \leq t \leq T}$ *is a martingale*

Theorem 4.3. (Discounted Feynman-Kac formula) Consider the stochastic differential equation

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

Let f be a Borel-measural function and let r be constant. Fix $T > 0$ and let $t \in [0, T]$ be given. Define the function

$$h(t, x) = \mathbb{E}^{t,x}\left[e^{-r(T-t)}f(X_T)\right] = \mathbb{E}\left[e^{-r(T-t)}f(X_T)|X_t = x\right]$$

Assume that $h(t, x) < \infty$ for all (t, x). Then $h(t, x)$ satisfies the partial differential equation

$$h_t(t, x) + \beta(t, x)h_x(t, x) + \frac{1}{2}\gamma^2(t, x)h_{xx}(t, x) = rh(t, x)$$

with the terminal condition $h(T, x) = f(x)$ for all x

Example 4.4. *Suppose that (S_t) satisfies the stochastic differential equation*

$$dS_t = rS_tdt + \sigma S_tdW_t$$

Let

$$f(t, x) = \mathbb{E}^{t,x}\left[h(S_t)\right]$$

Then $f(t, x)$ satisfies the partial differential equation

$$f_t(t, x) + rxf_x(t, x) + \frac{1}{2}\sigma^2x^2f_{xx}(t, x) = rf(t, x)$$

with terminal condition $h(T, x) = f(x)$. This equation could be solved numerically.

Theorem 4.5. (2-dimensional Feynman-Kac formula) Let $W_t = (W_t^1, W_t^2)$ be a 2-dimensional Brownian motion. Consider two stochastic differential equations

$$dX_t = \beta_1(t, X_t, Y_t)dt + \gamma_{11}(t, X_t, Y_t)dW_t^1 + \gamma_{12}(t, X_t, Y_t)dW_t^2$$

$$dX_t = \beta_2(t, X_t, Y_t)dt + \gamma_{21}(t, X_t, Y_t)dW_t^1 + \gamma_{22}(t, X_t, Y_t)dW_t^2$$

Let $f(x, y)$ be Borel measurable function, define

$$g(t, x, y) = \mathbb{E}^{t,x,y}\left[f(X_T, Y_T)\right]$$

$$h(t, x, y) = \mathbb{E}^{t,x,y}\left[e^{-r(T-t)}f(X_T, Y_T)\right]$$

Then g and h satisfy the partial differential equations

$$g_t + \beta_1 g_x + \beta_2 g_y + \frac{1}{2}(\gamma_{11}^2 + \gamma_{12}^2)g_{xx} + (\gamma_{11}\gamma_{12} + \gamma_{21}\gamma_{22})g_{xy} + \frac{1}{2}(\gamma_{21}^2 + \gamma_{22}^2)g_{yy}$$

$$h_t + \beta_1 h_x + \beta_2 h_y + \frac{1}{2}(\gamma_{11}^2 + \gamma_{12}^2)h_{xx} + (\gamma_{11}\gamma_{12} + \gamma_{21}\gamma_{22})h_{xy} + \frac{1}{2}(\gamma_{21}^2 + \gamma_{22}^2)h_{yy}$$

with terminal conditions

$$g(T, x, y) = h(T, x, y) = f(x, y)$$

for all x,y

5 Black-Scholes Economy and Black-Scholes Formula

There are two assets: a risky stock S and riskless bond B: These assets are driven by the SDEs

$$dS_t = \mu S_t dt + \sigma S_t dW_t$$

$$dB_t = r_t B_t dt$$

The time zero value of the bond is $B_0 = 1$ and that of the stock is S_0. The model is valid under certain market assumptions.
Let $C(t, S_t)$ be the European call price. With Ito lemma, it must satisfy

$$dC = \left[C_t + \mu S C_S + \frac{1}{2}\sigma^2 S^2 C_{SS}\right]dt + \sigma S C_S dW$$

Let V denote the dollar value of a portfolio with x shares of stock and y dollars of deposit

$$V = xS + y$$

If this is self-financing portfolio,i.e., $y = V - xS$, then

$$dV = xdS + yrdt$$

Let $x = C_s$. Then $V - C$ is a risk-free portfolio, and we have

$$d(V - C) = \left[rV - rSC_S - C_t - \frac{1}{2}\sigma^2 S^2 C_{SS}\right]dt = r(V - C)dt$$

Now we have the Black-Scholes PDE

$$C_t + rSC_S + \frac{1}{2}\sigma^2 S^2 C_{SS} - rC = 0$$

with $V(T) = C(T, S_T) = max(S_T - K, 0)$
Solving $C_t + rSC_S + \frac{1}{2}\sigma^2 S^2 C_{SS} - rC = 0$ gives us,

$$C(t, S_t) = S\mathcal{N}(d_1) - Ke^{-r(T-t)}\mathcal{N}(d_2)$$

where $\mathcal{N}(\cdot)$ is the standard normal CDF and

$$d_1 = \frac{log(\frac{S_t}{K}) + (r + \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}$$

$$d_2 = d_1 - \sigma\sqrt{T-t}$$

And this is the famous *Black-Scholes Formula.* Besides, Put-call parity is

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

And European put-option formula is

$$P(t, S) = Ke^{-r(T-t)}\mathcal{N}(-d_2) - S\mathcal{N}(-d_1)$$

Remark 5.1. *By Ito Lemma the value S_t of a derivative written on the stock follows the diffusion, and here we do not assume the functional form of V_t.*

$$dV_t = \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}dS + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}(dS)^2$$

$$= (\frac{\partial V}{\partial t} + \mu S_t \frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial S^2})dt + (\sigma S_t \frac{\partial V}{\partial S})dW_t$$