NUMERICAL ANALYSIS FOR FINANCIAL MATHEMATICS

~~Hirbod Assa~~ [1]

University of Liverpool

Ehsan. azmoodeh.

Office hour: Tue: 11:00~12:00

Tue: 17:00~18:00

office: Math Building 301C.

[1]Mailing address: Institute for Financial and Actuarial Mathematics. University of Liverpool. UK. Email: assa@liverpool.ac.uk

Contents

Introduction

This is the lecture-notes of "*Numerical methods in financial mathematics*". In this lecture-notes we care about three major issues in numerical methods:

1) Problem relevance;
2) Accuracy;
3) Time feasibility.

Our problems are all relevant in the context of economics, finance and insurance. Examples include application in

- Economics:
  - Demand and supply,
  - Optimal allocation,
- Finance:
  - Pricing options (trees),
  - Pricing options (Monte-Carlo),
- Insurance:
  - Insurance prices,
  - Risk capital management,

and many other interesting applications. The lecture-notes consists of six chapters. The first four chapters are respectively on basic numerical methods, numerical methods in option pricing, Monte-Carlo simulation methods in finance and Numerical methods for ordinary and stochastic differential equations. The fifth chapter consists of MATLAB codes used in the first four chapters, and chapter six consists of the solution of the selected exercises.

Chapters one to four, constitute the major content of the lecture-notes. Each chapter has multiple examples, and a good number of exercises. There are parts written in gray that are unnecessary for the exam, and are beneficial for students who want to acquire further knowledge. This lecture-notes comes with a handful number of videos called "*Numerical methods on-the-go*". Furthermore, a series of videos for MATLAB programming with a short note is provided to students.

Few points need to be explained. First of all all the figures, examples, numerical results, codes and exercises are original. Second, I would like to thank my former PhD student, Simon (Meng) Wang and José de Jesús Rocha Salazar for their help in preparing this document and the lecture-notes slides.

I hope this lecture-notes is useful.

CHAPTER 1

Basic numerical methods

1.1. Root finding

Root of a function is a point at which the value of the function is equal to zero. The root finding problem is one of the most important problems in all areas of science. This, of course, includes application in finance and insurance. For instance, one can find the equilibrium price or find the optimal portfolio. A root of a function on Euclidean spaces is introduced as follows.

欧式空间 $n \le 3$ 推广到有限n维空间.

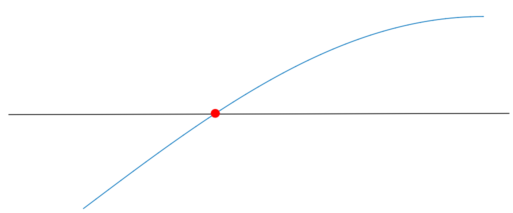
c.

Definition 1. Let $f : R^n \to R^m$, $\overrightarrow{r} \in R^n$ is a root if $f(\overrightarrow{r}) = \mathbf{0} \in R^m$.

let $f: C \subseteq R \to R$ be a function

domain

A point $c \in C$ is called a root of f if $f(c)=0$



E.g.: $f(x) = e^{-\tanh(\ln)}$

这种…

Figure 1.1.1. The root of a function is where it meets the horizon line.

Let us consider the following interesting example in pricing.

Example 2. Let us consider you are working in a manufacturing company (for instance home appliances manufacturing), and you are entitled to estimate the price of a product in the market. For that, you need to estimate the so-called demand and supply functions (you really do not need to know how they can be found). Based on some market research you know the demand function, which is a decreasing function, is equal to $d(x) = \frac{1}{x+1}$ and the supply function, which is an increasing function, is given by $s(x) = \exp(x) - 1$. In order to find the equilibrium price you need to find c such that $d(c) = s(c)$. In other words, if we introduce $f(x) = d(x) - s(x)$ then we need to find the answer of the following problem : $f(c) = 0$. Therefore, finding the root of f can give you the market price. As one can see maybe finding the solution to $\frac{1}{x+1} - \exp(x) + 1 = 0$ is not possible, therefore we need some numerical methods to find a "good enough" estimation of c.

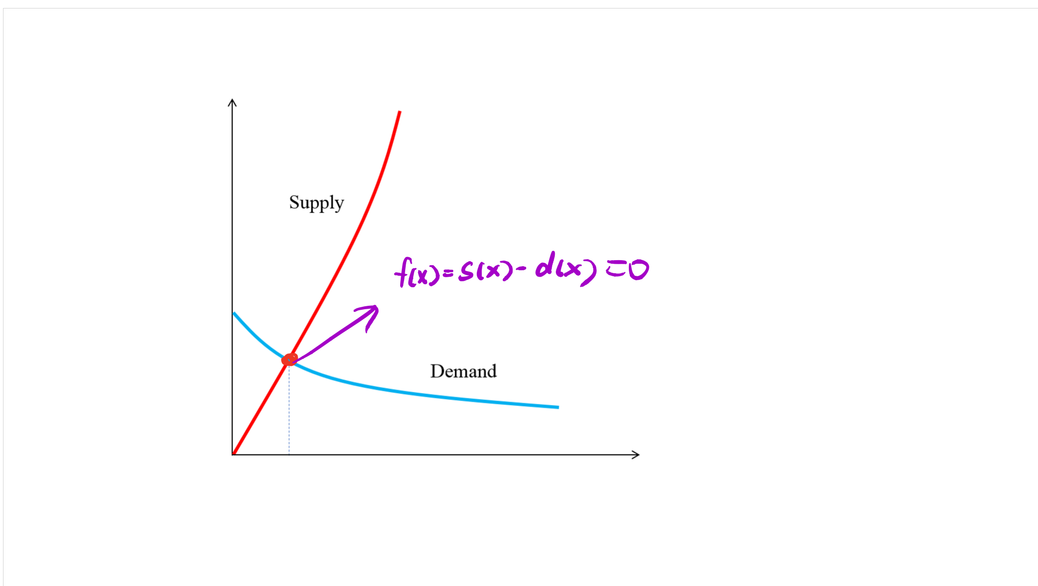


FIGURE 1.1.2. The intersection of the blue curve, the demand $d(x) = \frac{1}{x+1}$, and the red curve, the supply $s(x) = \exp(x) - 1$.

Among many methods we will discuss three

(1) Bisection method;

(2) Fixed point method;

(3) Newton method;

(4) Secant method.

Intermediate Value Theorem: Let $f: [a,b] \to R$ be a continous function. For $\forall u$: $\min\{f(a), f(b)\} \le u \le \max\{f(a), f(b)\}$. Then $\exists x \in (a,b)$; $f(x) = u$.

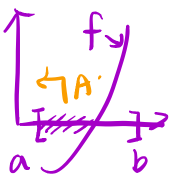
$f([a,b])$ is an interval.

$\sup(A) = x$ means that $a \le x$, $\forall a \in A$.

1.1.1. Bisection method. The bisection method is based on a very famous theorem stated below: (Binary Search Method) (二分法).

THEOREM 3. *Let $f : [a, b] \to R$ be a continuous function such that $f(a)f(b) < 0$. Then, $\exists r \in (a, b)$; $f(r) = 0$.*

Proof: Assume that $f(a) < 0$ and $f(b) > 0$. Define $A = \{x \in [a,b] ; f(x) < 0\}$ has to be negative. $a \in A \neq \Phi$. set $x = \sup A$. Now $f(x) = 0$ must be



This theorem gives us an idea to find the root by dividing the interval $[a, b]$ to two sub-intervals and choose the one in which we can again check the conditions of the theorem. step ① Find $a, b \in R$ so that $f(a)f(b) < 0$

Therefore, we need to find the middle point, step ② set $c_0 = \frac{a+b}{2}$ and check three case that can happen. compute $f(c_0)$

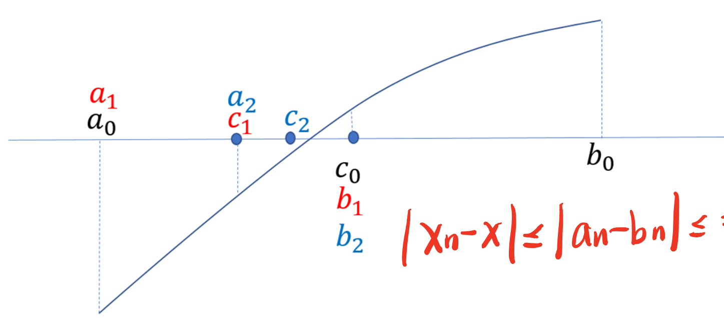
step ③ (1) If $f(c) = 0 \to$ root is found.

step ④ (2) If $f(a)f(c) < 0$ then $r \in (a, c)$, set $b = c$,

(3) If $f(b)f(c) < 0$ then $r \in (c, b)$ set $a = c$,

That means, except the root is found in 1, we update the interval to one with half length; i.e., if 2 happens, (a, c) is our new interval and if 3 happens, (c, b) is our new interval.

gives you $(x_n ; n \ge 1)$ c_n such that $x_n \to x$ (true root) as $n \to \infty$



$|x_n - x| \le |a_n - b_n| \le \frac{1}{2}|a_{n-1} - b_{n-1}| \le \dots \le (\frac{1}{2})^n |a_0 - b_0| < \varepsilon$ (given)

$n \to \infty$, $\to 0$.

FIGURE 1.1.3. The starting point in the bisection algorithm.

$2^n \ge \frac{b-a}{\varepsilon} \iff n \log 2 \ge \log(\frac{b-a}{\varepsilon})$

$\iff \boxed{n \ge \frac{\log(\frac{b-a}{\varepsilon})}{\log 2}}$ 要记.

minimal value

you need to iterate.

Here is an algorithm designed based on the idea explained above.

Algorithm 1 The bisection algorithm.

(1) First set, $a_0 = a$, $b_0 = b$
(2) Set

$$c_i = \frac{a_i + b_i}{2}$$

(3) if $f(c_i) = 0 \to r = c_i \to$ stop
(4) if $f(c_i)f(a_i) < 0 \to a_{i+1} = a_i, b_{i+1} = c_i$
(5) if $f(c_i)f(b_i) < 0 \to a_{i+1} = c_i, b_{i+1} = b_i$
(6) If $|f(a_i)| < \delta \to r = c_{i+1} \to$ stop, else, go to 2.

stopping criteria: ① $E_n < \varepsilon$ (given) ② given i. (the number of iteration)

We give a sketch of the proof since it can help us to see how fast this algorithm can converge. Indeed, we show that c_i converges to r. In iteration number i, we see that $r \in [a_i, b_i]$. On the other hand, in the next iteration we have

$$c_i = \frac{a_i + b_i}{2}.$$

Therefore,

$$\begin{aligned} |r - c_i| &= \left| r - \frac{a_i + b_i}{2} \right| \\ &\le \max \left\{ \left| \frac{a_i + b_i}{2} - a_i \right|, \left| b_i - \frac{a_i + b_i}{2} \right| \right\} = \frac{|a_i - b_i|}{2} \end{aligned}$$

But,

$$|a_i - b_i| = \frac{1}{2}|a_{i-1} - b_{i-1}| = \ldots = \frac{1}{2^i}|a_0 - b_0|$$

$\to 0$ as $i \to \infty$

Hence, we get

$$|r - c_i| \le \frac{1}{2^i}|a_0 - b_0| \longrightarrow 0, \text{ as } i \to \infty.$$

So one can see that, first the fact that ① c_i converges to r and second how fast ② it goes to the real root.

Example 4. Let us consider the function $f(x) = \frac{1}{x+1} - \exp(x) + 1 = 0$. Starting points are $a = 0$ and $b = 0.6$. To reach accuracy of $\delta = 0.0000001$, we need to iterate 22 times, and the estimated value is $r =$

Example 1: Consider $f(x) = x^3 + 4x^2 - 10$.

$f(1) = -5 < 0$ and $f(2) = 14 > 0$

$\Rightarrow$ Bolzano's theorem tells us that there is at least one root $[1, 2]$

- $x_0 = \frac{1+2}{2} = 1.5$ $\quad f(a) f(x_0) < 0$. $\quad x_0 \to b_1$

  $f(1.5) = 2.375 > 0 \Rightarrow [1, 1.5]$

- $x_1 = \frac{1+1.5}{2} = 1.25$ $\quad f(x_1) f(b_1) < 0$.

  $f(1.25) = -1.99 < 0 \Rightarrow [1.25, 1.5]$

- $x_2 = \frac{1.25+1.5}{2} = 1.375$

  $\vdots$

0.5086. In the following we show 10 iteration of the bisection method where you can see how the intervals change from right to left.

	$f(a)$	$f(c)$	$f(b)$
$n=1$	1	**0.419371962**	**-0.1971188**
$n=2$	0.419371962	**0.121342987**	**-0.1971188**
$n=3$	**0.121342987**	**-0.034721143**	-0.1971188
$n=4$	0.121342987	**0.044028381**	**-0.034721143**
$n=5$	0.044028381	**0.004842368**	**-0.034721143**
$n=6$	**0.004842368**	**-0.014891038**	-0.034721143
$n=7$	**0.004842368**	**-0.005012393**	-0.014891038
$n=8$	**0.004842368**	**-8.20452E-05**	-0.005012393
$n=9$	0.004842368	**0.002380901**	**-8.20452E-05**
$n=10$	0.002380901	**0.001149613**	**-8.20452E-05**

TABLE 1. 10 iteration of the bisection method

式子的值等于点本身

1.1.2. Fixed point method. Fixed point of a function is a point at which a function's value is equal to itself. Finding a fixed point of a function is a challenging problem. However, we use this problem to find the root of a function.

First, let us introduce the fixed point of a function.

DEFINITION 5. Let $g : C \subseteq R^n \to C \subseteq R^n$, be a function, $\vec{x} \in C$ is fixed point if $g(\vec{x}) = \vec{x}$. Root finding $\longleftrightarrow$ Fixed-Point

f $\quad g = x - \lambda f(x)$; $\lambda \neq 0$

Consider a function $f : C \to C$. To find the root of the function $f : C \to C$ when $C \subseteq R^n$, one can find the fixed point of the following function, $f(x)=0$ $\quad g(x_0)=x_0$

$$g(\vec{x}) = \vec{x} - \lambda f(\vec{x})$$

for a $\lambda \neq 0$.

Observe, that $(\lambda \neq 0)$

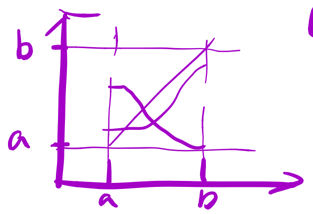
$$g(\vec{x}) = \vec{x} \Leftrightarrow \vec{x} = \vec{x} - \lambda f(\vec{x}) \Leftrightarrow f(\vec{x}) = 0.$$

Therefore, finding root of a function is equivalent to finding the fixed point of another function. As a result, we focus our attention to finding

Thm: (sufficient theorem)

(a) let $g: [a,b] \xrightarrow{con} R$, such that $g([a,b]) \leq [a,b]$

Then g has at least one fixed point in $[a,b]$



(b) If in addition, g is differentiable on (a,b), and $\exists\, q \in (0,1)$ such that $|g'(x)| \leq q$ $\forall x \in (a,b)$. Then the fixed point is unique.

prove (a): $f(a)=a$ or $f(b)=b$, we are done. since $f([a,b]) \leq [a,b]$,

$f(a) > a$, and $f(b) < b$

Define $h(x) = g(x) - x$.

$h(a) > 0$, $h(b) < 0$. So Bolzano's thm tells us that h has at least one root.

(b) 用 intermediate value thm. 证明

Fixed Point Method:

$g: [a,b] \to R$ given

① starting point, select (elegantly) $x_0 \in [a,b]$

② $x_n = g(x_{n-1}), n \geq 1$ $(*)$

Thm: let g be cont. and $g[a,b] \subset [a,b]$, then $(x_n)_{n\geq 1}$ given by $(*)$ coverge to a fixed point

Proof: since $|x_n| \leq |b-a|$, $\forall n \geq 1$ is a bounded sequence

so $x_n \to x \in [a,b]$ as $n \to \infty$.

$$x = \lim_{n\to\infty} x_n = \lim_{n\to\infty} g(x_{n-1}) = g\left(\lim_{n\to\infty} x_{n-1}\right) = g(x)$$

a fixed point of a function. In general, it is not always possible to find the fixed point of a function; that is why we need to impose some conditions and rely on some theorems to find a fixed point.

THEOREM 6. *Assume $g : C \subseteq R \to C \subseteq R$ is a function such that,*

$$\exists 0 < q < 1, \forall x, y \in C, |g(x) - g(y)| < q|x - y|.$$

(*)

Then g has a fixed point in C.

To find the fixed point of a function that satisfies the conditions of the theorem we use the following iterative methods:

define $x_n = g(x_{n-1})$, $\forall n \geq 1$

Then, we claim that $(x_n ; n \geq 1)$ converge to a point $x \in C$, and in fact $f(x) = x$

$$x_{k+1} = g(x_k), x_0 \in C.$$

We give a sketch of the proof since it can help us to see how one can generate an algorithm and how fast this algorithm can converge. Let x_0 be an arbitrary member in $C \subseteq R$. By iteration introduce,

$$x_{k+1} = g(x_k),$$

for $k = 0, 1, 2, \ldots$ We claim that x_k converges to a fixed point. First observe,

$$
\begin{aligned}
|x_{n+1} - x_n| &\overset{\text{def}}{=} |g(x_n) - g(x_{n-1})| \overset{(*)}{\leq} q|x_n - x_{n-1}| \quad \text{(decreasing)} \\
&= q|g(x_{n-1}) - g(x_{n-2})| \leq q^2|x_{n-1} - x_{n-2}| \\
&= q^2|g(x_{n-2}) - g(x_{n-3})| \leq q^3|x_{n-2} - x_{n-3}| \\
&\ldots. \\
&\leq q^n|x_1 - x_0| \to 0 \text{ as } n \to \infty
\end{aligned}
$$

} repeating.

(因为 $q<1$) $\Rightarrow$ enough to say it's converge.

That means (Take $m > n \geq 1$)

$$|x_n - x_0| = \left|\sum_{i=1}^{n}(x_i - x_{i-1})\right|$$

$$\leq \sum_{i=1}^{n}|x_i - x_{i-1}|$$

$$\leq \sum_{i=1}^{n}q^{i-1}|x_1 - x_0| \leq \frac{1}{1-q}.$$

$|x_m - x_n| = \left|\sum_{k=n+1}^{m}(x_k - x_{k-1})\right|$

$\leq \sum_{k=n+1}^{m}|x_k - x_{k-1}|$

$\leq \sum_{k=n+1}^{m} q^{k-1}|x_1 - x_0| = |x_1 - x_0| q^n(1 + q + q^2 \ldots + q^{m-n-1}) \to 0$ as $n, m \to \infty$.

$|x_n - x_{n-1}| \to 0 \Rightarrow$ cauchy sequence

then $\exists x \in C$

$x_n \to x$ as $n \to \infty$ ∠ take $m > n \geq 1$

enough.

Next, $|x - x_n| = \lim_{m\to\infty} |x_m - x_n| \leq q^n|x_1 - x_0| \lim_{m\to\infty}\left(\sum_{k=0}^{m-n-1} q^k\right) \leq q^n|x_1 - x_0| \left(\sum_{k=0}^{\infty} q^k\right) = \frac{q^n}{1-q}|x_1 - x_0|$

$\frac{1}{1-q}$ upper bound.

S_n is controlled

It means $\{x_n\}_n$ is a bounded sequence so it has a convergent subsequent that we denote it by x. Since $x_n - x_0 = \sum_{i=1}^{n}(x_i - x_{i-1})$ then $x - x_0 = \sum_{i=1}^{\infty}(x_i - x_{i-1})$, so, we have

$$\begin{aligned}
|x_n - x| &= \left|\sum_{i=1}^{n}(x_i - x_{i-1}) - \sum_{i=1}^{\infty}(x_i - x_{i-1})\right| \\
&= \left|\sum_{i=n+1}^{\infty}(x_i - x_{i-1})\right| \\
&\leq \sum_{i=n+1}^{\infty}|x_i - x_{i-1}| \\
&\leq \sum_{i=n+1}^{\infty} q^{i-1}|x_1 - x_0| \leq \frac{q^n}{1-q}.
\end{aligned}$$

This means x_n converges to x and at with rate q^n. Finally, since g is continuous,

$$g(x) = g(\lim_n x_n) = \lim_n g(x_n) = \lim_n x_{n+1} = x.$$

Now, the question is how practically we can use this method. For that we need to make sure a $0 < q < 1$ exists. Let us state the following theorem:

THEOREM 7. (b) *Let $g : [a,b] \to [a,b]$ be a continuous function which is differentiable and that,*

$$\sup_{x\in(a,b)} |g'(x)| < 1$$

Then one can get $q = \sup_{x\in(a,b)} |g'(x)|$.

Let us prove this theorem. Let $q = \sup_{x\in(a,b)} |g'(x)|$. Then by mean-value theorem for any $x, y \in (a,b), \exists \mu_{xy} \in (x,y)$ such that $g(x) - g(y) = g'(\mu_{xy})(x-y)$. Therefore, $|g(x) - g(y)| = |g'(\mu_{xy})||x-y| \leq q|x-y|$.

Let us go back to root finding problem of f. As we told root of f can be found by finding the fixed point of $g(x) = x - \lambda f(x)$. According to previous theorems, it is enough to find supremum of g and see whether we can regulate the amount of λ such that $\sup_{x\in(a,b)} |g'(x)| < 1$. So assume

f is continuously differentiable, then,

$$g'(x) = (x - \lambda f(x))' = 1 - \lambda f'(x)$$

Let λ be such that $1-\lambda f'(r)$ be too small. In that case g is a contraction around root r and therefore, has a fixed point by iteration.

This means in the best situation we can choose,

$$\lambda = \frac{1}{f'(r)}$$

which gives $g'(r) = 0$. Because g is continuously differentiable, for an interval I_r around r, we have

$$\forall x \in I_r \Longrightarrow |g'(x) - g'(r)| < q < 1 \Rightarrow |g'(x)| < q < 1.$$

So we have the following algorithm.

Algorithm 2 A fixed point algorithm.

(1) Choose λ a number close to $\frac{1}{f'(r)}$.
(2) Choose, $x_0 \in C$
(3) Set $x_{i+1} = x_i - \lambda f(x_i)$.
(4) If $|f\,(x_{i+1})\,| < \delta \to r = x_{i+1} \to$ stop, else, go to 3.

Let us consider the function $f(x) = \frac{1}{x+1} - \exp(x) + 1 = 0$. Let's take $\lambda = -1/2$ (which is close to $1/f'(0.5)$; see the example in the section on bisection method for approximation of r). So we have to find the fixed point of $g(x) = x + \frac{1}{2(x+1)} - \frac{1}{2}\exp(x) + \frac{1}{2}$. Starting from $x_0 = 0.6$, it takes only 6 iterations to get $r = 0.5086$ with accuracy $\delta = 0.0000001$.

	x_n
$n = 1$	0.6
$n = 2$	0.5014406
$n = 3$	0.508905042
$n = 4$	0.508536771
$n = 5$	0.508555642
$n = 6$	0.508554677

TABLE 2. Six iteration for a fixed point method to find the root.

The Newton Method

$f: [a,b] \xrightarrow{C^2} \mathbb{R}$, we are interated in the root if let assume c is a root of f and c_0 be an initial approximation. ($|c-c_0| \approx 0$), and moverover, assume that $f'(c_0) \neq 0$, Using Taylor Formula, we can write

$$\cancel{\underset{=0}{f(c)}} = f(c_0) + f'(c_0)(c-c_0) + \frac{f''(\xi(c,c_0))}{2}\cancel{(c-c_0)^2} \text{ (很小可略)}$$

$$0 \approx f(c_0) + f'(c_0)(c-c_0)$$

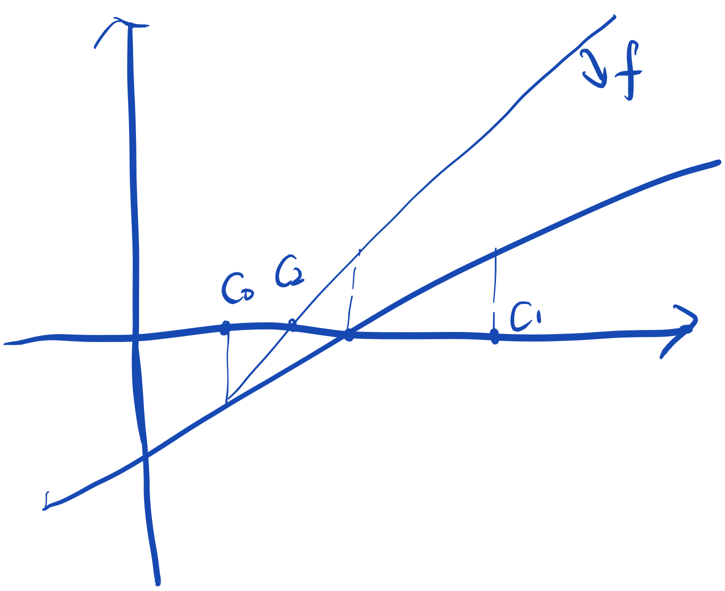
$$c \approx c_0 - \frac{f(c_0)}{f'(c_0)} \Rightarrow \text{is a fixed point method following the pattern}$$

① start with a "good" initial approximation x_0.

② $x_n = x_{n-1} - \dfrac{f(x_{n-1})}{f'(x_{n-1})}$; $\forall n \geq 1$

$$g(x) = x - \frac{f(x)}{f'(x)}$$

Geometric Intereption



$(x_k, f(x_k))$

$l: y = f'(x_k)(x-x_k) - f(x_k)$

put $y=0$. $\Rightarrow x = x_k - \dfrac{f(x_k)}{f'(x_k)}$

1.1.3. Newton's method. But the problem is we do not know $f'(r)$. That is why we can modify the method by replacing constant number λ with $f'(x)$. That means we can look at,

$$\boxed{g(x) = x - \frac{1}{f'(x)} f(x) = x - \frac{f(x)}{f'(x)},}$$

which yields the Newton's method as follows. Let us assume f is two times continuously differentiable function such that $f'(r) \neq 0$. Then, starting from a "good" initial point x_0 then, the sequence

$$x_{k+1} = g\left(x_k\right) = x_k - \frac{f(x_k)}{f'(x_k)}.$$

is a sequence converging to r.

THEOREM 8. *Assume f is a two times continuously differentiable function such that $f'(r) \neq 0$. Then there is a $\delta > 0$ such that if $p_0 \in (-\delta + r,\ \delta + r)$ then*

$$\underline{x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}}$$

is a sequence converging to r.

PROOF. Since $f(r) = 0$, $f'(r) \neq 0$ then for $q < 1$ there exists $\delta > 0$ such that $|\frac{f(x)f''(x)}{f'(x)^2}| \leq q < 1$. Now let us look at $g(x) := x - \frac{f(x)}{f'(x)}$. It can be seen easily that

$$\begin{aligned} g'(x) &= 1 - \frac{f'(x)^2 - f''(x)f(x)}{f'(x)^2} \\ &= 1 - 1 + \frac{f''(x)f(x)}{f'(x)^2} = \frac{f''(x)f(x)}{f'(x)^2} \end{aligned}$$

This indicates $|g'(x)| = |\frac{f''(x)f(x)}{f'(x)^2}| \leq q < 1$ for any $x \in (-\delta + r,\ \delta + r)$. So we have $x_{k+1} = g(x_k)$ will converge to a fixed point of g. So there is a $\widetilde{x}$ such that $g(\widetilde{x}) = \widetilde{x}$ and $\widetilde{x} - \frac{f(\widetilde{x})}{f'(\widetilde{x})} = \widetilde{x} \Longrightarrow f(\widetilde{x}) = 0 \Longrightarrow \widetilde{x} = r$.

□

We can also look at the Newton's method from a different perspective. Let us look at Taylor's expansion

$x \sim x_0$中的一个数.

$$f(x) = f(x_0) + (x - x_0)f'(x_0) + (x - x_0)^2 \frac{f''(\mu)}{2} \quad for\ \mu \in (x, x_0)$$

For root r one can have

$$0 = f(r) = f(x_0) + (r - x_0)f'(x_0) + (r - x_0)^2 \frac{f''(\mu)}{2}$$

Note $(r - x_0)^2 \frac{f''(\mu)}{2}$ is too small to zero if x_0 is close to r. Therefore, we have $0 \approx f(x_0) + (r - x_0)f'(x_0) \Rightarrow r \approx x_0 - \frac{f(x_0)}{f'(x_0)}$. And a good estimation for r is

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Geometrical perspective of Newton's method. Let us consider the line that is tangent to the function f and passes through $(x_k, f(x_k))$. Then, the line is given by $y = f'(x_k)(x - x_k) - f(x_k)$. Let us look at the point that hits the x axes i.e., when $y = 0$. We find out that

$$y = 0 \Rightarrow x = x_k - \frac{f(x_k)}{f'(x_k)} = x_{k+1}.$$

It means if we draw the tangent line tangent to $(x_k, f(x_k))$, the line hits the x axes at x_{k+1} which is the next point in the iteration. That is how one can find the next point in the Newton method.

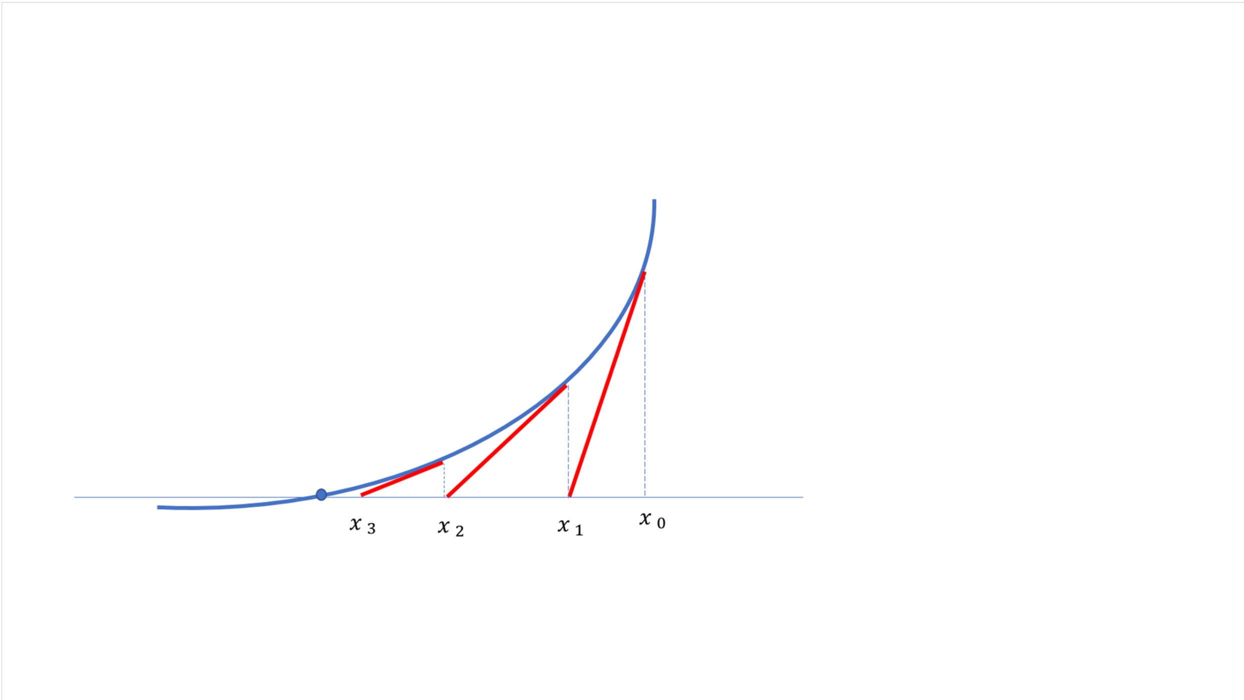


FIGURE 1.1.4. Newton's method.

Algorithm 3 The algorithm for Newton's method.

---

(1) First set, $x_0 \in C$
(2) Set
$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}.$$
(3) If $|f(x_{i+1})| < \delta \rightarrow r = x_{i+1} \rightarrow$ stop, else, go to 2.

EXAMPLE 9. Let us consider the function $f(x) = \frac{1}{x+1} - \exp(x) + 1 = 0$. Then $f'(x) = -\frac{1}{(1+x)^2} - \exp(x)$. Then the Newton method is to find the fixed point of $g(x) = x + \frac{\frac{1}{x+1} - \exp(x) + 1}{\frac{1}{(1+x)^2} + \exp(x)}$. Starting from $x_0 = 0.6$, it takes only 4 iterations to get $r = 0.5086$ with accuracy $\delta = 0.0000001$.

Example : $x^3+4x^2-10=0$ has a only root on $[1,2]$

$f(x)=0 \Leftrightarrow g(x) = x - \lambda f(x)$, $\lambda \neq 0$.

$4x^2 = 10 - x^3 \Rightarrow x = \pm\frac{1}{2}(10-x^3)^{\frac{1}{2}}$

(a) $x = g_1(x)$ where $g_1(x) = \frac{1}{2}(10-x^3)^{\frac{1}{2}}$.

$\Rightarrow |g_2'(x)| = |\sqrt{10}(-\frac{1}{2})(4+x)^{-\frac{3}{2}}|$, $x \in [1,2]$

(b) $x^2(x+4)-10=0$

$\Rightarrow x = \left(\frac{10}{4+x}\right)^{\frac{1}{2}} \Rightarrow g_2(x) = \left(\frac{10}{4+x}\right)^{\frac{1}{2}}$.

(c) $x = g_3(x) = x - \frac{x^3+4x^2-10}{3x^2+8x}$

$g_1(1)=6,\ g_1(2)=-12,$

$g([1,2]) \not\subseteq [1,2]$

$= \frac{\sqrt{10}}{2} \times 5^{-\frac{3}{2}} \leq 0.15 < 1.$

Newton Method.

n	(a)	(b)
0	1.5	1.5
1	-0.625	1.348
		⋮
3 / 4	1.03×10^8.	1.365.

Actural root IS 1.365230013

	x_n
$n=1$	0.6
$n=2$	0.510916573
$n=3$	0.508556161
$n=4$	0.508554724

Table 3. Four iteration of the Newton's method.

1.1.4. Secant method. Sometimes we do not have f' or it is not computationally easy to find it. In that case we can use an approximation. We know,

$$f'(x) = \lim_{h\to 0} \frac{f(x+h)-f(x)}{h}$$

Therefore, if in an iterative method x_k, x_{k-1} are very close then,

$$f'(x_k) \approx \frac{f(x_k)-f(x_{k-1})}{x_k - x_{k-1}}$$

So from Newton's method,

Newton →

$x_k = x_{k-1} - \frac{f(x_k)}{f'(x_{k-1})} \overset{(*)}{\Rightarrow} x_k = x_{k-1} - \frac{x_{k-1}-x_{k-2}}{f(x_{k-1})-f(x_{k-2})} f(x_{k-1})$

$$x_{k+1} = x_k - \frac{f(x_k)}{\frac{f(x_k)-f(x_{k-1})}{x_k-x_{k-1}}}$$

$f'(x_{k-1}) = \lim_{x\to x_{k-1}} \frac{f(x)-f(x_{k-1})}{x-x_{k-1}}$ $x = x_{k-2}$

or

$(*) \approx \frac{f(x_{k-2})-f(x_{k-1})}{x_{k-2}-x_{k-1}}.$

$$x_{k+1} = x_k - (x_k - x_{k-1})\frac{f(x_k)}{f(x_k)-f(x_{k-1})}.$$

Geometrical perspective of Secant's method. Let us consider the line that is passing through $(x_{k-1}, f(x_{k-1}))$ and $(x_k, f(x_k))$. Then, the line is given by $y = \frac{f(x_k)-f(x_{k-1})}{x_k-x_{k-1}}(x-x_k) - f(x_k)$. Let us look at the point that this line hits the x axes; i.e., $y=0$. Then, we find out that

$$y = 0 \Rightarrow x = x_k - (x_k - x_{k-1})\frac{f(x_k)}{f(x_k)-f(x_{k-1})} = x_{k+1}.$$

It means if we draw the line passing through $(x_{k-1}, f(x_{k-1}))$ and $(x_k, f(x_k))$, the line hits the x axes at x_{k+1} which is the next point in the iteration. That is how one can find the next point in the Secant method.

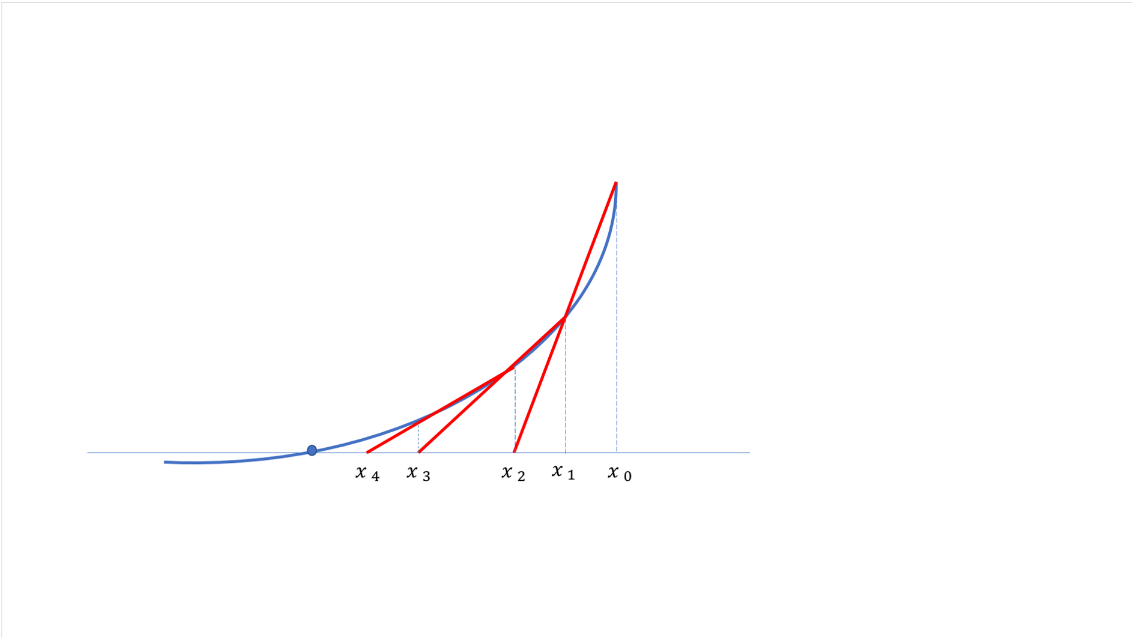


FIGURE 1.1.5. Secant method.

Algorithm 4 The algorithm for the Secant method.

---

(1) First set, $x_0, x_1 \in C$

(2) Set
$$x_{i+1} = x_k - (x_k - x_{k-1})\frac{f(x_k)}{f(x_k) - f(x_{k-1})}.$$

(3) If $|f(x_{i+1})| < \delta \to r = x_{i+1} \to$ stop, else, go to 2.

EXAMPLE 10. Let us consider the function $f(x) = \frac{1}{x+1} - \exp(x) + 1 = 0$. Starting from $x_0 = 0.7$, and $x_1 = 0.6$, it takes only 5 iterations to get $r = 0.5086$ with accuracy $\delta = 0.0000001$.

	x_n
$n = 1$	0.6
$n = 2$	0.513695273
$n = 3$	0.508682149
$n = 4$	0.508554893
$n = 5$	0.508554724

TABLE 4. Five iteration of the Secant method.

Optimization in Finance
Central Framework
min $f(x)$ = $f: \mathbb{R}^n \to \mathbb{R}$

Free-Derivative optimization Method.

Derivative optimization Method.
⇓
Gradient Method

$x \in C$ objective function

set of constraint feasible region

The Nelder-Mead Method
very general Method
No assumption on smoothness of f

Gradient Method
required some smoothness Assumption of f.

1.2. Minimization

1.2.1. Introduction. There are many problems in finance and insurance that the main objective is to find an optimal solution. For instance, we can minimize the risk, maximize the profit or the utility. Let me first introduce the minimum of a function:

Definition 11. Let $f : C \subseteq R^n \to R$. Then the minimum of f happens at a point $m \in C$ where $f(m) \leq f(x), \forall x \in C$.

Let us look at the following example:

Example 12. Let us consider a portfolio management problem as follows

$$\begin{cases} \min \left\{a_1^2 + a_2^2 + a_3^2 + a_4^2\right\} \\ a_1 + a_2 + a_3 + a_4 = 1 \\ a_1 + 2a_2 + 2a_3 + a_4 = 2 \end{cases}.$$

From second and third relation we have $a_1 = -a_4$ and $a_2 = 1 - a_3$. If we put in the first relation we get

$$\min_{a_1, a_2} \left\{a_1^2 + (1 - a_1)^2 + a_2^2 + (1 - a_2)^2\right\}.$$

So one can see in order to find the optimal portfolio we have to minimize a function in R^2.

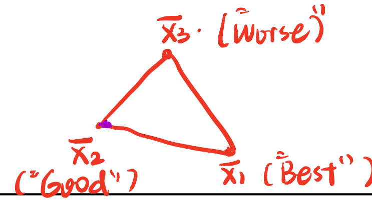
Among all the methods, in this lecture notes we introduce the following ones

(1) Nelder-Mead method
(2) Gradient method (or steepest decent)

1.2.2. Nelder-Mead method. Nelder-Mead method is a method to find the minimum of a function $f : C \subset R^n \to R$. The main advantage of this method is that the function f does not need to be too smooth. However, this method does not converge very fast and also can only give the local minimums. In this lecture we only explain this method for the case when $n = 2$, i.e., for a function $f : C \subseteq R^2 \to R$. For R^n the method is similar but a little bit complicated.

This method starts from an initial simplex (a triangle in R^2), then by applying an operator, from a list of four operators, at each iteration we update the simplex until we get close enough to the minimum. The idea is to search a better update by replacing the point at which the function gets the largest value by getting closer to the two points at which the function takes the smallest values.

The method is explained in the following algorithm:



Algorithm 5 The Nelder-Mead algorithm.

(1) Define a simplex $\vec{x_i} = (x_i, y_i) \in R^2$, $i = 1, 2, 3$. Initial Point $= \bar{x}_1, \bar{x}_2, \bar{x}_3$

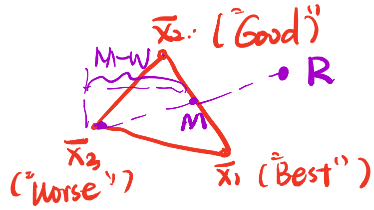
(2) Sort the value and by rearranging subscripts order them as follows,

② $z_i = f(x_i, y_i)$ $z_1 \le z_2 \le z_3$. Relabel them as $B = z_1$ $G = z_2$ and

worse $W = z_3$. So, $f(B) \le f(G) \le f(W) \le f(R)$ Best Good.

(3) Find mid point of G and B,
③
$$M = \frac{B+G}{2} = \left(\frac{x_1+x_2}{2}, \frac{y_1+y_2}{2}\right)$$

Find the reflected point towards M,



$$R = M + (M - W) = 2M - W$$

Compute $f(R)$.

(4) Reflection. If $f(B) < f(R) < f(G)$. Then R is good enough and we can replace W with R and go to step one.

Expension point

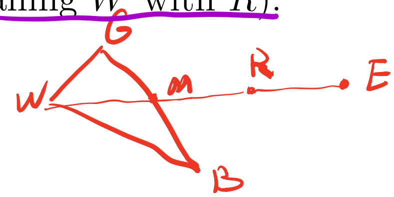
(5) Expansion. If $f(R) \le f(B)$. Define $E = 2R - M$. then we moved in right direction and we do expansion. Check:

(a) If $f(E) < f(R)$ replace W with E (and not R).

(b) Else keep the previous replacement (meaning W with R).

4 operation.

(6) Contraction. If $f(G) \le f(R)$, let



$$C_1 = \frac{R+M}{2}$$
$$C_2 = \frac{M+W}{2}.$$

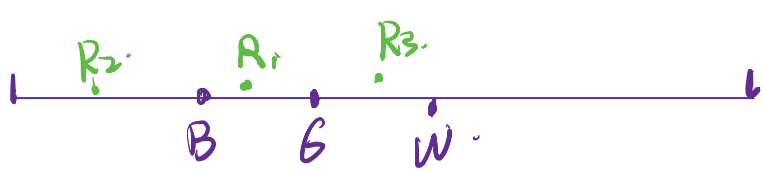
If $f(C_1) \le f(W)$ or $f(C_2) \le f(W)$, replace W with C_i where C_i is the best.

(7) Shrink. If non of the steps above happens let

$$S = \frac{B+W}{2}$$

Replace W with S. Replace G with M.

(8) If $\frac{|f(B)-f(W)|}{\|B-W\|} < \delta$ Stop, else go to 2.



- Mid-point.

Remark: when to stop?: ① Dimeter of triangle $< \delta$ ← given.
② Area of triangle $< \delta$
③ $|f(B) - f(W)| < \delta$.

Example: $f(x,y) = 4(x-5)^2 + 6(y-6)^2$. $|B-W|$

start with initial points

$\vec{x}_1 = (8,9)$, $\vec{x}_2 = (10,11)$, $\vec{x}_3 = (8,11)$

$\underbrace{f(\vec{x}_1) = 90}_{B}$, $\underbrace{f(\vec{x}_2) = 250}_{W}$, $\underbrace{f(\vec{x}_3) = 186}_{G}$.

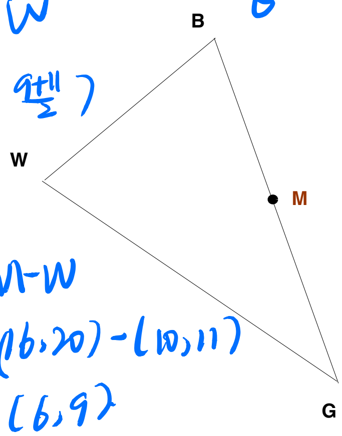
$M = \dfrac{B+G}{2} = \left(\dfrac{8+8}{2}, \dfrac{9+11}{2}\right)$

$= (8,10)$

Refection $R = 2M - W$

$= (16,20) - (10,11)$

$= (6,9)$



$f(R) = f(6,9) = 58$.

FIGURE 1.2.1. Finding mid point.

$f(R) < f(B) \Rightarrow$ Expension

$E = 2R - M$

$= 2(6,9) - (8,10)$

$= (4,8)$

• Reflection.

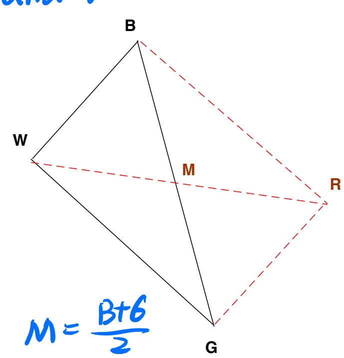
$f(E) = 4 + 24 = 28 < f(R) = 58$.

so, Expension is accepted, and we replace W with E.

$\vec{x}_1{}' = (8,11) \Rightarrow f(\vec{x}_1{}') = 186$

$\vec{x}_2{}' = (8,9) \Rightarrow f(\vec{x}_2{}') = 90$

$\vec{x}_3{}' = (4,8) \Rightarrow f(\vec{x}_3{}') = 28$.



$\vec{x}_3{}' = B$, $\vec{x}_2{}' = G$, $\vec{x}_1{}' = W$, $M = \dfrac{B+G}{2}$

$= (6,8.5)$

$R = 2M - W = 2(6,8.5) - (8,11)$

$= (4,6)$

FIGURE 1.2.2. Finding reflection point.

$f(R) = f(4,6) = 4 < f(B) = 28$.

we are in right direction, so we do expension

$E = 2R - M = 2(4,6) - (6,8.5) = (2,3.5)$

• Expansion.

$f(E) = f(2,3.5) = 73.5 > f(R)$

Expansion Point E is not accepted, an we just replace W with R.

New Triangle: $\vec{x}_1{}^2 = (4,6)$ $\vec{x}_2{}^2 = (8,9)$ $\vec{x}_3{}^2 = (4,8)$

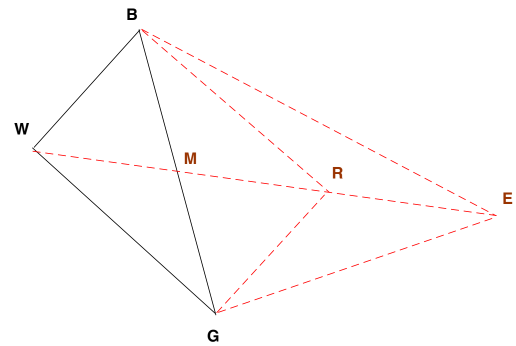


FIGURE 1.2.3. Finding extension point.

- Contraction.

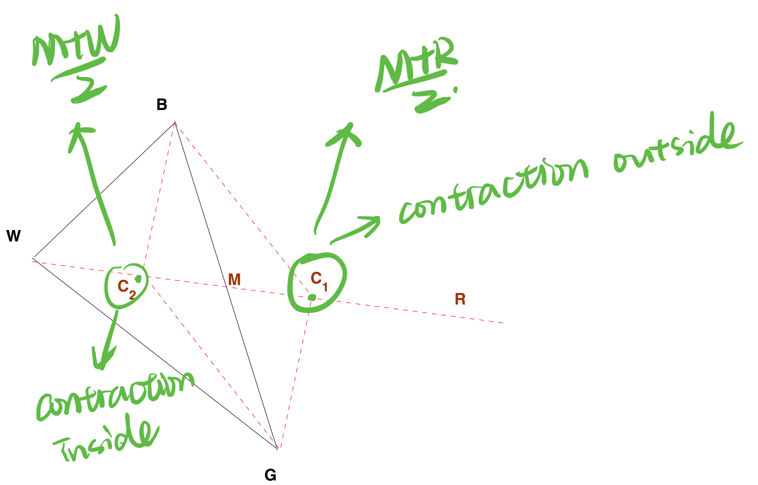


FIGURE 1.2.4. Finding contraction points.

- Shrink.

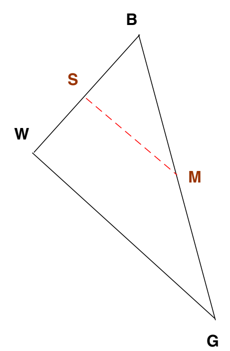


FIGURE 1.2.5. Finding shrink points.

EXAMPLE 13. Consider that we want to minimize the function $f(x, y) = x^2 + y^2$ with Nelder-Mead method.

1) Starting with simplex $\overrightarrow{x_1} = (1, 0)$, $\overrightarrow{x_2} = (0, 1)$ and $\overrightarrow{x_3} = (4/5, 4/5)$, we complete the first iteration of Nelder-Mead method, and determine the new simplex and the values on each vertex.

2) We have to order $x_1, x_2$and x_3.Observe that

$$f(\overrightarrow{x_1}) = 1^2 + 0^2 = 1, f(\overrightarrow{x_2}) = 0^2 + 1^2 = 1,$$

$$f(\overrightarrow{x_3}) = \left(\frac{4}{5}\right)^2 + \left(\frac{4}{5}\right)^2 = \frac{32}{25}.$$

So it is clear that $W = \overrightarrow{x_3}$, but since the value of f on $\overrightarrow{x_1}, \overrightarrow{x_2}$ are equal, one can take one B and one G. So we take $B = \overrightarrow{x_1}$ and $G = \overrightarrow{x_2}$.

3) Now, we find mid point and the reflection point:

$$M = \frac{B+G}{2} = \left(\frac{1+0}{2}, \frac{0+1}{2}\right) = \left(\frac{1}{2}, \frac{1}{2}\right).$$

$$R = 2M - W = 2\left(\frac{1}{2}, \frac{1}{2}\right) - \left(\frac{4}{5}, \frac{4}{5}\right) = \left(\frac{1}{5}, \frac{1}{5}\right).$$

4) Reflection. Now we have to check $f(B) < f(R) < f(G)$. But $f(B) = 1 < f(R) = \frac{2}{25} < f(G) = 1$, which does not hold. So we have to proceed to the next step.

5) Expansion. Now we need to check $f(R) \leq f(B)$. So $f(R) = \frac{2}{25} < f(B) = 1$ which holds and we need to find the expansion point

$$E = 2R - M = 2\left(\frac{1}{5}, \frac{1}{5}\right) - \left(\frac{1}{2}, \frac{1}{2}\right) = \left(-\frac{1}{10}, -\frac{1}{10}\right).$$

Now we have to check $f(E) = \frac{2}{100} < f(R) = \frac{2}{25}$, which holds, that means we have to replace W with E and the new simplex is $(1,0), (0,1)$ and $\left(-\frac{1}{10}, -\frac{1}{10}\right)$.

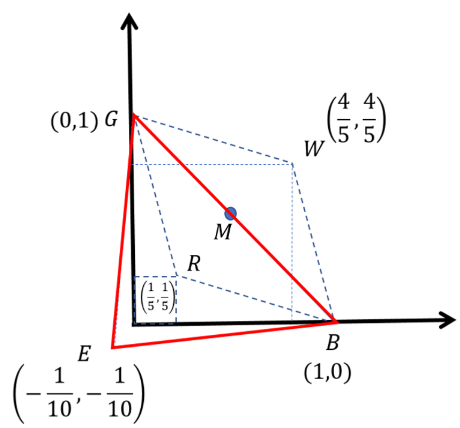


FIGURE 1.2.6. Nelder-Mead method.

1.2.3. Gradient method (Steepest descant). When one wants to move towards minimum of a function (think of a surface), one needs to get the direction with steepest decent. Mathematically, this direction is nothing but the direction of minus gradient $(-\nabla f)$. This method can take us step by step to the minimum.

Derivative Method in Optimization.

$f: C \subseteq R^n \to \mathbb{R}$, $\min f(x)$

Assumption: $\nabla f(\vec{x}) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots \frac{\partial f}{\partial x_n}\right)$

② $\|\nabla f(\vec{x}_0)\|^2 = \sum_{k=1}^{n} \left(\frac{\partial f}{\partial \vec{x}_k}(\vec{x}_0)\right)^2$.

③ $\vec{x}_{n+1} = \vec{x}_n - \gamma \frac{\nabla f(\vec{x}_n)}{\|\nabla f(\vec{x}_n)\|}$

[choosing γ]

$\ldots \frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3})$

① Start with some initial point $\vec{x}_0 \in R^n$.

② $\vec{x}_1 = x_0 - \gamma \frac{\nabla f(\vec{x}_0)}{\|\nabla f(\vec{x}_0)\|}$ → γ.

① γ is given to you
② If is not given to you, then you need to find a "good" one

第一

Let $f : C \subseteq R^n \to R$ be a real function with continuous derivatives. Denote,

法一 $\varphi(\gamma) := f(\vec{x}_0 - \gamma \frac{\nabla f(\vec{x}_0)}{\|\nabla f(\vec{x}_0)\|})$, $\varphi'(\gamma) = 0 \Rightarrow \gamma$. ← select γ which makes φ minimal

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\right).$$

法二 ③ If f is convex and L-Lipschitz then any choice $0 < \gamma < \frac{2}{L}$ is working

$|f(x) - f(y)| \le L|x - y|$

As explained above, in order to approach minimum we have to move towards $-\nabla f(\vec{x})$ or simply,

$$S_0(\vec{x}) = \frac{-\nabla f(\vec{x})}{||\nabla f(\vec{x})||},$$

$\sup_x |f'(x)| \le L$

$\approx \frac{f(x) - f(y)}{x - y}$.

where,

$$||\nabla f(\vec{x})|| = \sqrt{\left(\frac{\partial f}{\partial x_1}\right)^2, \ldots, \left(\frac{\partial f}{\partial x_n}\right)^2}.$$

$f: R^n \to R$, $\sup_x \|\nabla f(\vec{x})\| \le L$

So, let's start from $\vec{x_0}$, for some $\gamma > 0$, in the next iteration we update to $\vec{x_1} = \vec{x_0} + \gamma S(\vec{x_0})$; this gives us a point which has moved towards $-\nabla f$, therefore it take f closer to the minimum,

When to stop?
① $\|\nabla f(\vec{x}_n)\| < \delta$ ← given. $\Rightarrow$ stop.
② $|f(\vec{x}_n)| < \delta \Rightarrow$ stop

$$f(\vec{x_0}) > f(\vec{x_1}).$$

Let $\vec{x_0}$ be a point close to minimum, $\vec{x_k}$ can be constructed iteratively as follows.

Algorithm 6 Algorithm for gradient method.

---

(1) Choose, $x_0 \in C$

(2) Set

$$S(\vec{x_i}) = \frac{\nabla f(\vec{x_i})}{||\nabla f(\vec{x_i})||_2} = \frac{\left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\right)}{\sqrt{\left(\frac{\partial f}{\partial x_1}\right)^2, \ldots, \left(\frac{\partial f}{\partial x_n}\right)^2}}$$

$$\vec{x}_{i+1} = \vec{x_i} + \gamma S(\vec{x_i}).$$

(3) If $||\nabla f(\vec{x_i})||_2 < \delta \to r = \vec{x}_{i+1} \to$ stop, else, go to step 2.

EXAMPLE 14. Consider that we want to minimize the function $f(x, y) = (x-1)^4 + (y-2)^2$ with gradient method. Then the gradient of function f is

$$\nabla f = \left(4(x-1)^3, 2(y-2)\right),$$

$\gamma = 0.9$.

$\vec{x}_{n+1} = \vec{x}_n - 0.9 \frac{\nabla f(\vec{x}_n)}{\|\nabla f(\vec{x}_n)\|}$, $n \ge 0$.

and

$$\|\nabla f\| = \sqrt{16\,(x-1)^6 + 4\,(y-2)^2}.$$

If we choose $\gamma = 0.1$, then the algorithm is to find iteratively (given that $\overrightarrow{x_i} = (x_i, y_i)$):

$$\overrightarrow{x}_{i+1} = \overrightarrow{x_i} - 0.1\frac{\left(4\,(x_i-1)^3\,, 2\,(y_i-2)\right)}{\sqrt{16\,(x_i-1)^6 + 4\,(y_i-2)^2}}.$$

1.3. Iterative methods for linear problems

$A = (a_{ij})_{1\le i,j\le n}$, $b \in R^n$.
$Ax = b$.

Let us assume that we want to solve a large system of equations. For instance, consider a large portfolio of n assets. The mean-variance minimization problem is given as follows

$$\min\left\{\frac{1}{2}\overrightarrow{x}A\overrightarrow{x}^T - \overrightarrow{x}b\right\},$$

where A is the symmetric, $n\times n$ positive definite co-variance matrix and b is the $n\times 1$ mean matrix. If we take the derivative and equalize it to zero we get the following system of linear equations

$$A\overrightarrow{x}^T = b.$$

Theoriotical solution.
How to calculate A^+ when n is very large?

Sometimes the number of the assets in the portfolio can be hundreds, i.e., $n > 100$. In a normal way, one need to find the inverse of A in order to find the solution $\overrightarrow{x}^T = A^{-1}b$. But it necessitates finding the determinate of a matrix whose dimension is at least 100×100. This is computationally impossible. That is why we need different methods to find the solution. In the following we will explain two iterative methods to approximate the solution in a feasible manner:

(1) Jacob method
(2) Gauss-Seidle method

The idea of both methods is to introduce a fixed point problem whose solution is equal to the solution of our system of linear equations and then find an approximation of the solution by using the iterative method we have discussed earlier in the section on fixed points.

1.3.1. *Jacobi method.* Let us assume we want to solve the following system of linear equations,

$$A\overrightarrow{x}^T = b$$

Assumption ①

where A is $n\times n$, $\overrightarrow{x}^T$ is $n\times 1$ and b is $n\times 1$. Assume A is symmetric and also that all member on diagonal are non zero, i.e.

$$a_{ij} = a_{ji}, a_{ii} \neq 0$$

②

Every matrix A can be decomposed as follows,

$$A = \underset{\text{lower}}{L} + \underset{\text{upper}}{U} + D$$

$$\begin{pmatrix} a_{11} & & a_{1n} \\ a_{21} & & \vdots \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} 0 & & \\ a_{21} & 0 & 0 \\ a_{31} & & \\ a_{n1} & \cdots & 0 \end{pmatrix} + \begin{pmatrix} 0 & a_{12} & \cdots & a_{1n} \\ a_{21} & 0 & & \vdots \\ a_{31} & & \ddots & \\ a_{n1} & & & 0 \end{pmatrix} + \begin{pmatrix} a_{11} & & 0 \\ & a_{22} & \\ 0 & & a_{nn} \end{pmatrix}$$

where L is lower triangular, U is upper triangular, and D is Diagonal.

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 6 & 7 & 8 \\ 3 & 7 & 2 & 3 \\ 4 & 8 & 3 & 8 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 3 & 7 & 0 & 0 \\ 4 & 8 & 3 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 2 & 3 & 4 \\ 0 & 0 & 7 & 8 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 8 \end{bmatrix}$$

So $A\overrightarrow{x}^T = b$ can be written again as follows,

$$\begin{aligned} A\overrightarrow{x}^T = b &\Rightarrow (L+U+D)\overrightarrow{x}^T = b \\ &\Rightarrow (L+U)\overrightarrow{x}^T + D\overrightarrow{x}^T = b \\ &\Rightarrow D\overrightarrow{x}^T = b - (L+U)\overrightarrow{x}^T. \end{aligned}$$

As a result we get

$$\overrightarrow{x}^T = -D^{-1}(L+U)\overrightarrow{x}^T + D^{-1}b \tag{1.3.1}$$

Since D is diagonal, D^{-1} is easily given by,

$$(D^{-1})_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ d_{ii}^{-1} & i = j \end{cases}$$

Introduce,

$$g\left(\overrightarrow{x}^T\right) = -D^{-1}(L+U)\overrightarrow{x}^T + D^{-1}b$$

It is clear that the solution to (1.3.1) is $\boxed{\text{a fixed point for } g.}$ Then, under appropriate conditions

$Ax=b$

$$\overrightarrow{x}^T_{k+1} = -D^{-1}(L+U)\overrightarrow{x}^T_k + D^{-1}b.$$

converges. One way to verify the convergence is to see whether A is diagonally dominant: $|a_{ij}| \geq \sum_{j\neq i} |a_{ij}|$.

Let us compute $\overrightarrow{x}^T_{k+1}$ in terms of $\overrightarrow{x}^T_k$, recall

① start with some Initial value $\vec{x}_0$
② $\vec{x}_{n+1} = g(\vec{x}_n) = D^{-1}b - D^{-1}(L+U)\vec{x}_n$.
} Jacobi Method.

$$L+U=\begin{bmatrix} 0 & a_{12} & a_{13} & ... & a_{1n} \\ a_{21} & 0 & a_{23} & ... & a_{2n} \\ a_{31} & a_{32} & 0 & ... & a_{3n} \\ ... & ... & ... & ... & ... \\ a_{n1} & a_{n2} & a_{n3} & ... & 0 \end{bmatrix} \text{ and } D^{-1}=\begin{bmatrix} a_{11}^{-1} & 0 & 0 & 0 & 0 \\ 0 & a_{22}^{-1} & 0 & 0 & 0 \\ 0 & 0 & a_{33}^{-1} & 0 & 0 \\ 0 & 0 & 0 & ... & 0 \\ 0 & 0 & 0 & 0 & a_{nn}^{-1} \end{bmatrix}.$$

This gives us

$$D^{-1}(L+U)=\begin{bmatrix} 0 & a_{12}/a_{11} & a_{13}/a_{11} & ... & a_{1n}/a_{11} \\ a_{21}/a_{22} & 0 & a_{23}/a_{22} & ... & a_{2n}/a_{22} \\ a_{31}/a_{33} & a_{32}/a_{33} & 0 & ... & a_{3n}/a_{33} \\ ... & ... & ... & ... & ... \\ a_{n1}/a_{nn} & a_{n2}/a_{nn} & a_{n3}/a_{nn} & ... & 0 \end{bmatrix}.$$

So we have

$$\begin{aligned} D^{-1}(L+U)x^k =& 0 + x_2^k\frac{a_{12}}{a_{11}} + x_3^k\frac{a_{13}}{a_{11}} + ... + x_n^k\frac{a_{1n}}{a_{11}} + \\ & x_1^k\frac{a_{21}}{a_{22}} + 0 + x_3^k\frac{a_{23}}{a_{22}} + ... + x_n^k\frac{a_{2n}}{a_{22}} + \\ & ... \end{aligned}$$

On the other hand we have $D^{-1}b = \left(\frac{b_1}{a_{11}}, \frac{b_2}{a_{22}}, ..., \frac{b_n}{a_{nn}}\right)^T$. And this give us

$$x_i^{k+1} = -\frac{1}{a_{ii}}\sum_{j\neq i} x_j^k a_{ij} + \frac{b_{i.}}{a_{ii}}.$$

Example 15. Let us assume that we have 100 assets in a portfolio numbered by 1,..., 100. Let us also assume the co-variance between asset i, j is given by $\exp\exp(-2|i-j|)$. If the vector of the means is given by $b = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$, then based on the discussions we had in the beginning of this section we need to solve the following system of equations

$$A\overrightarrow{x}^T = b,$$

where $A = [\exp(-2|i-j|)]_{i,j}$ and $b = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$. But note that $D_{ii} = A_{ii} = \exp(-2|i-i|) = 1$. So $D = I$. So using the Jacobi method we consider the following function

$$g\left(\overrightarrow{x}^T\right) = -D^{-1}(L+U)\overrightarrow{x}^T + D^{-1}b = -(A-I)\overrightarrow{x}^T + b.$$

After 16 iteration one can reach the accuracy of of 0.0000001. For illustration sake we have reduced the number to $n = 3$, and in the following table show the 12 iteration needed to reach accuracy of 0.0000001.

$n = 1$	1	1	1
$n = 2$	0.846349078	0.729329434	0.846349078
$n = 3$	0.88579457	0.770918216	0.88579457
$n = 4$	0.879443671	0.760241482	0.879443671
$n = 5$	0.881004931	0.761960483	0.881004931
$n = 6$	0.880743694	0.761537896	0.880743694
$n = 7$	0.88080567	0.761608605	0.88080567
$n = 8$	0.880794965	0.76159183	0.880794965
$n = 9$	0.880797431	0.761594728	0.880797431
$n = 10$	0.880796994	0.76159406	0.880796994
$n = 11$	0.880797092	0.761594179	0.880797092

FIGURE 1.3.1. Twelve iteration for Jacobi method.

1.3.2. Gauss-Seidel method. Let us consider the same linear problem $A\overrightarrow{x}^T = b$ in the last section. One can come up with a different fixed point problem as follows:

$$\begin{aligned} A\overrightarrow{x}^T = b &\Rightarrow (L+U+D)\overrightarrow{x}^T = b \\ &\Rightarrow (L+D)\overrightarrow{x}^T = -U\overrightarrow{x}^T + b \\ &\Rightarrow \overrightarrow{x}^T = -(L+D)^{-1}U\overrightarrow{x}^T + (L+D)^{-1}b. \end{aligned}$$

Let $g\left(\overrightarrow{x}^T\right) = -(L+D)^{-1}U\overrightarrow{x}^T + (L+D)^{-1}b$. So, the iteration method is given as follows

$$\overrightarrow{x}^T_{k+1} = -(L+D)^{-1}U\overrightarrow{x}_k^T + (L+D)^{-1}b.$$

$\overrightarrow{x}^T = -D^{-1}(L+U)\overrightarrow{x}^T + D^{-1}b$

$\frac{1}{a_{ii}}\left(b_i - \sum_{j\neq i} a_{ij} x_{kj}\right)$

Gauss-Seidel is more likely to converge.

Following the same logic as above, we can have

$$x_i^{k+1} = -\frac{1}{a_{ii}} \sum_{j=i+1}^{n} x_j^k a_{ij} - \frac{1}{a_{ii}} \sum_{j=1}^{i-1} x_j^{k+1} a_{ij} + \frac{b_{i.}}{a_{ii}}.$$

Note here, we also use iteration $k+1$ to compute x_i^{k+1} $(j \leq i-1)$.

Example 16. Let us assume that we have 100 assets in a portfolio numbered by 1,..., 100. Let us also assume the covariance between asset i, j is given by $\exp(-2|i-j|)$. If the vector of the means is given by $b = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$, then based on the discussions we had in the beginning of this section we need to solve the following system of equations

$$A\overrightarrow{x}^T = b,$$

where $A = [\exp(-2|i-j|)]_{i,j}$ and $b = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$. But note that $D_{ii} = A_{ii} = \exp(-2|i-i|) = 1$. So $D = I$. So using the Jcoby mthod we consider the following function

$$g\left(\overrightarrow{x}^T\right) = -(L+D)^{-1}U\overrightarrow{x}^T + (L+D)^{-1}b.$$

After only 10 iteration one can reach the accuracy of of 0.0000001. For illustration sake we have reduced the number to $n = 3$, and in the following table show the only 7 iteration needed to reach accuracy of 0.0000001.

$n=1$	1	1	1
$n=2$	0.882980356	0.750123825	0.846349078
$n=3$	0.880837066	0.761094007	0.88230943
$n=4$	0.88079781	0.761579683	0.880864033
$n=5$	0.880797091	0.761593794	0.880799023
$n=6$	0.880797078	0.761594148	0.880797127

Figure 1.3.2. Seven iteration for Gauss-Seidel method.

1.4. Interpolation methods

~~Let us consider that we have two variables that are related through a function. In other words, if we have two parameters x and y, we~~

接高斯

Ex: $12x_1+3x_2-5x_3=1$

$x_1+5x_2+3x_3=28$

$3x_1+7x_2+13x_3=76.$

Applied G-S Method with initial apr.

$X^{(0)}=(1,0,1)$

$$A=\begin{pmatrix} 12 & 3 & -5 \\ 1 & 5 & 3 \\ 3 & 7 & 13 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} 1 \\ 28 \\ 76 \end{pmatrix}$$

$12 > 3+|-5|$, $5 > 1+3$, $13 > 3+7$

So it converge to true solution.

A is strictly diagonally dominant.

E_1 $x_1=(1-3x_2+5x_3)/12$

E_2 $x_2=\dfrac{28-x_1-3x_3}{5}$

E_3 $x_3=\dfrac{76-3x_1-7x_2}{13}$

$X^{(0)}=(1,0,1)$

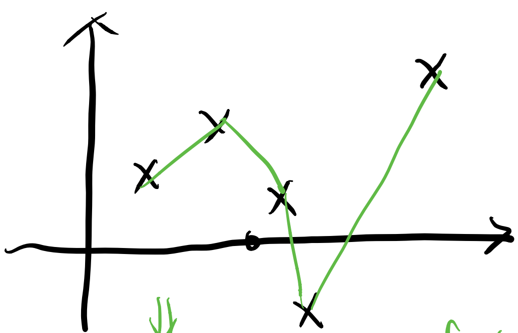
$X^{(1)}=(\frac{1}{3},4.9,3.09)$ ⇐ Apply Initial Value

$=(x_1^{(1)}, x_2^{(2)}, x_3^{(3)}).$ → $\dfrac{76-\frac{1}{3}\times 3-7\times 4.9}{13}$

更新为 $x_1^{(1)}$ ← $\dfrac{28-\frac{1}{3}-3}{5}$

★ Jacobi 的 $x_1^{(1)}, x_2^{(2)}, x_3^{(3)}$ 只基于 $X^{(0)}$,
而 G-S 的 $X^{(1)}$ 基于最新更新的 $x_1^{(1)}$...

Interpolation: Data $(x_i, f(x_i))$ $i=0,1,\cdots n.$

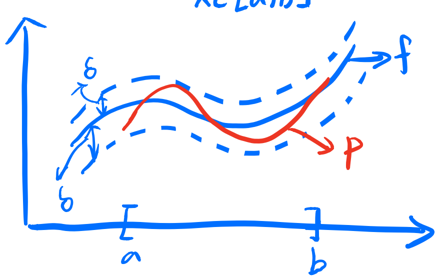


each straight line is first degree polynomial.

Recall: Thm [Weierstrass Approximation Theorem]

Let $f:[a,b]\xrightarrow{\text{Cont.}}\mathbb{R}$. Then $\forall\varepsilon>0$, $\exists$ a polynomial P

$$\|f-p\|=\sup_{x\in[a,b]}|f(x)-P(x)|<\varepsilon.$$



Piecewize-Linear Interpolation is simple to apply, but we can do much better

⇓

Lagrange Polynomial interpolation.

$(x_0, f(x_0))$ $\quad$ $(x_1, f(x_1))$

L_0 $\quad$ L_1

$L_0(x_0)=1$ $\quad$ $L_1(x_0)=0$

$L_0(x_1)=0$ $\quad$ $L_1(x_1)=1$

$L_0(x)=\dfrac{(x-x_1)}{x_0-x_1}$ 由 $L_0(x_1)=0$ 确定；由 $L_0(x_0)=1$ 确定 $\quad$ $L_1(x)=\dfrac{(x-x_0)}{(x_1-x_0)}$

We are looking for a polynomial P with propotion that $f(x_i)=P(x_i)$, $i=0,1$.

$P(x)=\underbrace{L_0(x)}_{\text{function}}f(x_0)+\underbrace{L_1(x)}_{}f(x_1)$ ⇒ Lagarian first degree...

接 1.4.1. number

1.4. Interpolation methods

Let us consider that we have two variables that are related through a function. In other words, if we have two parameters x and y, we know that there must be a function f, where $y = f(x)$. However, in the real word, one can make just few observations from x and y, say, $(x_1, y_1), \ldots, (x_n, y_n)$. With this few observation we are unable to retrieve the whole function f. For that reason we need to construct a "reasonable" approximation of this function. Even though in general one cannot say what is the most reasonable way to do this, there are mathematical ways to construct a "reasonable approximation".

Let us discuss the following example. In a market we are looking for a demand function of a product. In reality the demand function follows the function $D(x) = \frac{1}{(1+x)^2}$ (which in the first place we assume that we do not know). But we observe only few points, $(1, 1/4), (2, 1/9)$ and $(3, 1/16)$. Since we do not know the real function $D(x) = \frac{1}{(1+x)^2}$, one can come up with different functions. For instance, one can propose the following piece-wise linear function as a candidate for D:

$$g(x) = \begin{cases} \left(\frac{1}{9} - \frac{1}{4}\right)(x-1) + \frac{1}{4}, & x \in [1,2] \\ \left(\frac{1}{16} - \frac{1}{9}\right)(x-2) + \frac{1}{9}, & x \in [2,3] \end{cases}.$$

Somebody else can come up with the following function

$$h(x) = \frac{13}{288}x^2 + -\frac{79}{288}x + \frac{23}{48}.$$

One can see that $g(1) = h(1) = 1/4$, $g(2) = h(2) = 1/9$ and $g(3) = h(3) = 1/16$, which shows both functions can replicate the values of the observations. But the main question is which function can give a more reasonable estimation of function $D(x) = \frac{1}{(1+x)^2}$ (which is unknown). In the following subsections we will discuss two different methods:

(1) Polynomial interpolation

(2) Cubic spline

1.4.1. Polynomial interpolation. Let us assume that we have values of a function f on $n+1$ points $x_0, x_1, \ldots, x_n \in [a, b]$. We want to find a polynomial of $\boxed{\text{degree } n}$ which equals to f at these $n+1$ points as an approximation. If we denote the polynomial with,

$$P_n(x) = a_n x^n + a_{n-1}x^{n-1}, \ldots + a_1 x + a_0,$$

In general: $(x_i, f(x_i)), i=0, \cdots n$ $\forall\ 0 \le k \le n$, define

Define $P(x) = L_{n,0}^{(x)} f(x_0) + L_{n,1}^{(x)} f(x_1) + \cdots + L_{n,n}^{(x)} f(x_n)$.

$P(x_i) = f(x_i), \ \forall\ i = 0, \cdots, n.$

→ n-th Lagrange Polynomial Interpolating, $\deg(P) = n$.

$L_{n,k}(x) = \prod_{\substack{i=0 \\ i\neq k}} \frac{(x-x_i)}{(x_k-x_i)}$,

degree ↑, better approximation. (taylor).

Then $L_{n,k}(x_i)=0$, $i \neq k$,

$L_{n,k}(x_k)=1 \Rightarrow \deg(L_{n,k}) = n$

$n+1-1$.

where,

$$P(x_k) = f(x_k), k = 0, 1, 2, ..., n.$$

One way is to solve $n+1$ equations and $n+1$ unknown as,

$$P_n(x_i) = a_n x_i^n + a_{n-1} x_i^{n-1} + ... + a_1 x_i + a_0 = f(x_i).$$

Or, we can write it differently

$$\begin{bmatrix} x_0^n & x_0^{n-1} & \cdots & x_0 & 1 \\ x_1^n & x_1^{n-1} & \cdots & x_1 & 1 \\ \vdots & & \ddots & & \\ x_n^n & x_n^{n-1} & \cdots & x_n & 1 \end{bmatrix} \begin{bmatrix} a_n \\ a_{n-1} \\ \vdots \\ \vdots \\ a_0 \end{bmatrix} = \begin{bmatrix} f(x_0) \\ f(x_1) \\ \vdots \\ \vdots \\ f(x_n) \end{bmatrix}.$$

As a result P_n is unique.

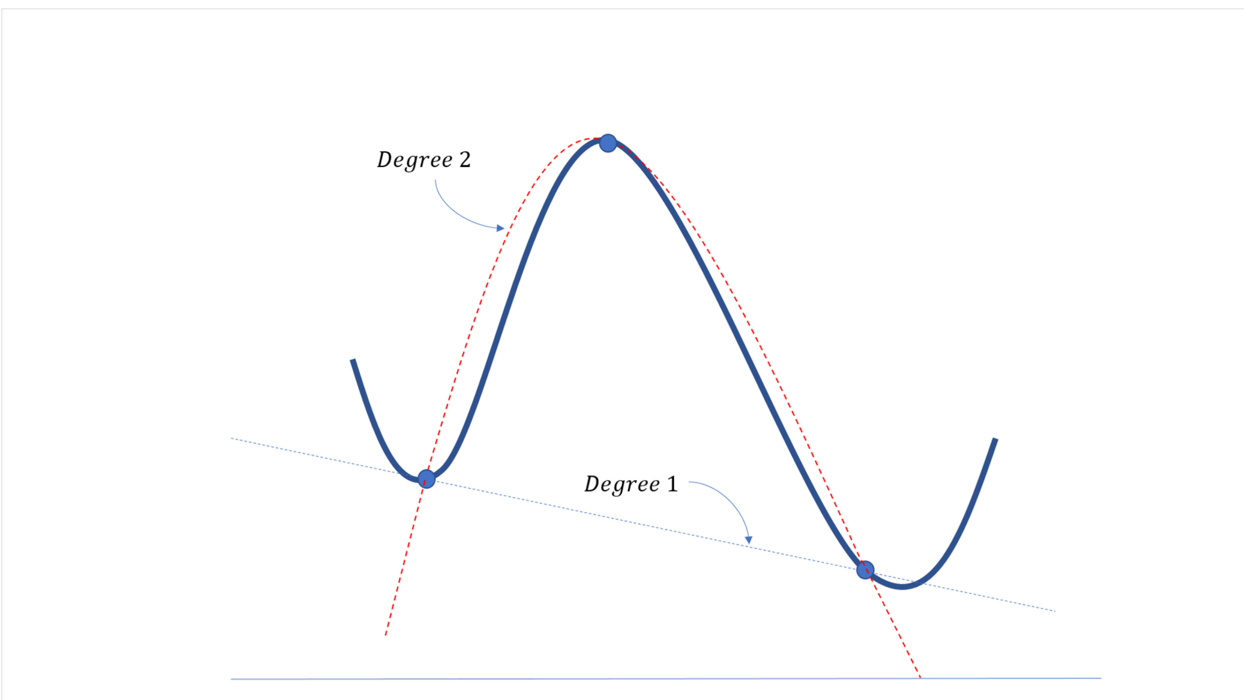


FIGURE 1.4.1. Degree 1 and degree 2 interpolation.

Finding the polynomial this way is very hard since we have to solve a system of linear equation whose components are growing exponentially; either can be very large or very small. Another way is to use Lagrangian polynomials. For $i = 0, \ldots, n$, introduce $n+1$ Lagrangian polynomials as follows

$$\boxed{L_i(x) = \frac{\Pi_{j\neq i}(x - x_j)}{\Pi_{j\neq i}(x_i - x_j)}.}$$

It is easy to see that,

3.

1. $L_i(x_i) = \frac{\Pi_{j\neq i}(x_i - x_j)}{\Pi_{j\neq i}(x_i - x_j)} = 1$
2. $L_i(x_j) = 0,\ j \neq i.$

2 $\quad f(x_i)\ \sum_0^n L_i(x_i)$

Let us introduce $\boxed{q(x) = \sum_0^n f(x_i)L_i(x).}$ Then, one can observe that

$$q(x_i) = 0\times f(x_0) + \ldots + 0\times f(x_{i-1}) + 1\times f(x_i) + 0\times f(x_{i+1}) + \ldots + 0\times f(x_n) = f(x_i).$$

So, q is a polynomial such that,

$$q(x_i) = f(x_i).$$

Therefore, q is what we are looking for. Error means difference between f and P_n. Let us introduce this difference as,

Error Analysis:

$x_i \in [a,b],\ i=0,1,\ldots,n.$

$$E_n(x) = f(x) - P_n(x).$$

n+1 distinct point

THEOREM 17. *If $f \in C^{n+1}[a,b]$ (i.e., $n+1$ times differentiable with continuous derivatives), then for any x there exists $c \in (a,b)$ such that,*

$$E_n(x) = \frac{(x - x_0)(x - x_1)\ldots(x - x_n)f^{n+1}(c)}{(n+1)!}.$$

的error就是这么大。

From this theorem one can derive the following estimations for the error term.

THEOREM 18. *If $\forall i, x_i - x_{i-1} = h$ for $n = 1, 2, 3, \ldots$ then,*

$$E_n(x) \leq C_n M_{n+1} h^{n+1},$$

where $M_n = \max_{x\in[a,b]} |f^n(x)|$. $\quad E_n(x) \leq \frac{M_{n+1}}{(n+1)!} h^{n+1}$

If $\max_{1\leq i\leq n} (x_i - x_{i-1}) \leq h$,

where $M_{n+1} = \sup_{x\in[a,b]} |f^{(n+1)}(x)|$

EXAMPLE 19. Let us consider that from a function $y = f(x)$, we only have access to three values $y = 1, 2, 3$, associated with $x = 1, 2, 4$, respectively. We want to find a curve which passes through these points. Assume we want to pass a polynomial through all these points. First of all the degree of the polynomial P_2 passing through these points is 2, since we

E.x.: $f(x) = \frac{1}{x}$, $x_0 = 2$, $x_1 = 2.75$, $x_2 = 4$ $\qquad f(3) \approx P_2(3) = 0.32955.$

$L_0(x) = \frac{(x-2.75)(x-4)}{(2-2.75)(2-4)} = \frac{2}{3}(x-2.75)(x-4)$

$L_1(x) = \frac{(x-2)(x-4)}{(2.75-2)(2.75-4)} = -\frac{16}{15}(x-2)(x-4)$

Next, we compute the maximum error, when we apply P_2 in approximating f

$L_2(x) = \frac{(x-2)(x-2.75)}{\ldots}$ $\qquad E = \frac{f'''(\xi(x))}{\ldots}(x-2)(x-2.75)(x-4)$

$L_2(x) = \frac{\cdots}{(4-2)(4-2.75)} = \frac{\cdots}{5}(x-2)(x-2.75)$ $E_2 = \frac{f'''(\xi(x))}{3!}(x-2)(x-2.75)(x-4)$

$P_0(x) = L_0(x)f(2) + L_1(x)f(2.75) + L_2(x)f(4)$; $f(2) \to \frac{1}{2}$, $f(2.75) \to \frac{4}{11}$, $f(4) \to \frac{1}{4}$. $f(x) = x^{-1}$, $f'(x) = -\frac{1}{x^2}$, $f''(x) = \frac{2}{x^3}$, $f'''(x) = 6x^{-4}$.

$|f'''(x)| = |-6x^{-4}| = 6 \times \frac{1}{16}$.

$g(x) = (x-2)(x-2.75)(x-4)$

$0 = g'(x) = (x-2.75)(x-4) + (x-2)(x-4) + (x-2)(x-2.75)$

$= \frac{1}{2}(3x-7)(2x-7)$

$\Rightarrow x = \frac{7}{3}, \frac{7}{2}$

$g(\frac{7}{3}) = \frac{25}{108}$ $g(\frac{7}{2}) = -\frac{9}{16}$

$|g(x)| \le \frac{9}{16}$

$\Rightarrow |E_2(x)| \le \frac{\frac{9}{16}}{16} = \frac{9}{16 \times 16}$

≈ 0.035

have access to three points. So we have to think of degree 2 Lagrangian polynomials. So we have

$$L_0(x) = \frac{(x-2)(x-4)}{(1-2)(1-4)} = \frac{(x-2)(x-4)}{3},$$

$$L_1(x) = \frac{(x-1)(x-4)}{(2-1)(2-4)} = -\frac{(x-1)(x-4)}{2},$$

$$L_2(x) = \frac{(x-1)(x-2)}{(4-1)(4-2)} = \frac{(x-1)(x-2)}{6}.$$

As a result the polynomial P_2 is

$$\begin{aligned} P_2(x) &= L_0(x) \times f(x_0) + L_1(x) \times f(x_1) + L_2(x) \times f(x_2) \\ &= \frac{(x-2)(x-4)}{3} - (x-1)(x-4) + \frac{(x-1)(x-2)}{2} \\ &= -\frac{1}{6}x^2 + \frac{3}{2}x - \frac{1}{3}. \end{aligned}$$

1.4.2. Piecewise cubic spline. Let us assume we have the value of a function f at $n+1$ points, $x_0, \ldots, x_n$ as $f(x_0), \ldots, f(x_n)$. Let us assume we want to approximate this function with piece-wise cubic polynomials smoothly i.e. $S(x_k) = f(x_k)$ $k = 0, \ldots, n$ and S' and S'' exist and are continuous. For that we associate a cubic polynomial S_k to each interval $[x_k, x_{k+1}]$ such that,

(b) (1) $S_k(x_k) = f(x_k)$

(c) (2) $S_k(x_{k+1}) = S_{k+1}(x_{k+1})$ (connect them), $k = 0, 1, \cdots n-1$

(d) (3) $S'_k(x_{k+1}) = S'_{k+1}(x_{k+1})$ (smoothly connect them) $k = 0, 1, \cdots n-2$. $2(n-1) = 2n-2$.

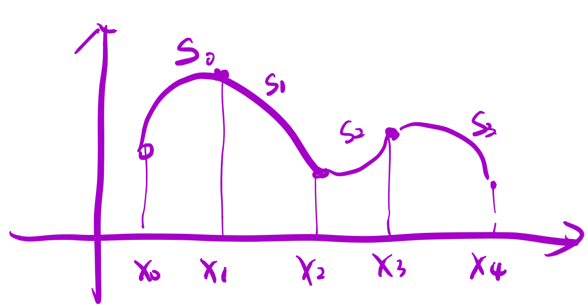
(e) (4) $S''_k(x_{k+1}) = S''_{k+1}(x_{k+1})$ (Smooth curvature)

$(x_i, f(x_i))$, $i = 0, \cdots n$.

P such that

$P(x_i) = f(x_i)$

$\forall i = 0, 1, \cdots n$.



Def: Given a function f on $[a,b]$, and a data set $\{(x_i, f(x_i))\}, i=0,1,\cdots n$. a cubic spline interpolating polynomial for f on $[a,b]$ is a piecewise-polynomial functions, denoted by S_k on the subinterval $[x_k, x_{k+1}]$; $k=0,\cdots n-1$, with the following properties:

(b) 接[illegible].

(f): boundary conditions ← Free / Natural Boundary Condition.

$S''(x_0) = X''(x_n) = 0$.

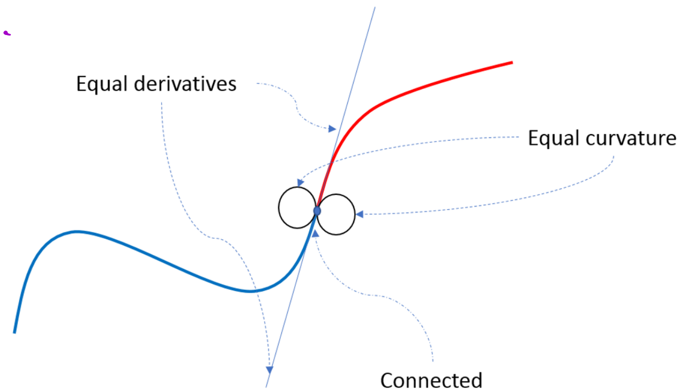


FIGURE 1.4.2. Schematic picture for cubic spline.

Since S_k are of degree 3 on$[x_k, x_{k+1}]$, then they can have the following form,

$$S_k(x) = S_{k,0} + S_{k,1}x + S_{k,2}x^2 + S_{k,3}x^3.$$

However, a smarter form that can save us lots of computation is the following form: 标记

$$S_k(x) = \frac{m_k}{6h_k}(x_{k+1} - x)^3 + \frac{m_{k+1}}{6h_k}(x - x_k)^3 + p_k(x_{k+1} - x) + q_k(x - x_k),$$

where$h_k = x_{k+1} - x_k$. ; $k=0,\cdots n-1$

To see why it is a smarter form, let us check rule number 4, meaning that let's take the second derivative: $S_k''(x) = \frac{m_k}{h_k}(x_{k+1} - x) - \frac{m_{k+1}}{h_k}(x - x_k)$. Now observe that

$$S_k''(x_{k+1}) = \frac{m_k}{h_k}(x_{k+1} - x) + \frac{m_{k+1}}{h_k}(x - x_k)\Big|_{x=x_{k+1}}$$

$$= \frac{m_{k+1}}{h_k}\underbrace{(x_{k+1} - x_k)}_{h_k} = \boxed{m_{k+1},}$$

(e)

and

$$\boxed{S''_{k+1}(x_{k+1})} = \frac{m_{k+1}}{h_{k+1}}(x_{k+2}-x) + \frac{m_{k+2}}{h_{k+1}}(x-x_{k+1})\Big|_{x=x_{k+1}}$$

$$= \frac{m_{k+1}}{h_{k+1}}\underbrace{(x_{k+2}-x_{k+1})}_{h_{k+1}} = \boxed{m_{k+1}.}$$

(f) ⇒ $m_0 = m_{k+1} = 0$.

As one can see the third condition already holds, so no need to be checked. Then, by following the rules 1 to 3 one can find m_k, p_k and q_k. First, combining rule number 1 and 2 gives us

apply (b).

$$S_k(x_k) = f(x_k) \Rightarrow \frac{m_k}{6}{h_k}^2 + p_k h_k = f(x_k)$$

$$\boxed{\Rightarrow p_k = \frac{f(x_k)}{h_k} - \frac{m_k}{6}h_k} \quad (*) \tag{1.4.1}$$

and

$$S_k(x_{k+1}) = f(x_{k+1}) \Rightarrow \frac{m_{k+1}}{6}{h_k}^2 + q_k h_k = f(x_{k+1}).$$

$$\Rightarrow q_k = \frac{f(x_{k+1})}{h_k} - \frac{m_{k+1}}{6}h_k. \tag{1.4.2}$$

Now us p_k, q_k in rule number 3, we get

$$\boxed{S'_k(x_{k+1}) = S'_{k+1}(x_{k+1})} \Rightarrow -\frac{m_k}{2h_k}(x_{k+1}-x)^2 + \frac{m_{k+1}}{2h_k}(x-x_k)^2 - p_k + q_k\Big|_{x=x_{k+1}}$$

condition (d)

$$= -\frac{m_{k+1}}{2h_{k+1}}(x_{k+2}-x)^2 + \frac{m_{k+2}}{2h_{k+1}}(x-x_{k+1})^2 - p_{k+1} + q_{k+1}\Big|_{x=x_{k+1}}$$

$$\Rightarrow \frac{m_{k+1}}{2}h_k - \frac{f(x_k)}{h_k} + \frac{m_k}{6}h_k + \frac{f(x_{k+1})}{h_k} - \frac{m_{k+1}}{6}h_k$$

$$= -\frac{m_{k+1}}{2}h_{k+1} - \frac{f(x_{k+1})}{h_{k+1}} + \frac{m_{k+1}}{6}h_{k+1} + \frac{f(x_{k+2})}{h_{k+1}} - \frac{m_{k+2}}{6}h_{k+1}$$

$$\Rightarrow \frac{m_{k+1}}{3}h_k + \frac{m_k}{6}h_k + \frac{f(x_{k+1})-f(x_k)}{h_k}$$

$$= -\frac{m_{k+1}}{3}h_{k+1} - \frac{m_{k+2}}{6}h_{k+1} + \frac{f(x_{k+2})-f(x_{k+1})}{h_{k+1}}$$

$$\Rightarrow 2m_{k+1}\left(h_k + h_{k+1}\right) + m_{k+2}h_{k+1} + m_k h_k$$

$$= 6\left(\frac{f(x_{k+2})-f(x_{k+1})}{h_{k+1}} - \frac{f(x_{k+1})-f(x_k)}{h_k}\right).$$

If we denote $d_{k+1} = \frac{f(x_{k+1})-f(x_k)}{h_k}$, then to find $m'_k s$ we have to solve

$$h_{k-1}m_{k-1} + 2(h_{k-1}+h_k)m_k + h_k m_{k+1} = 6(d_k - d_{k-1}),$$

☆ (***)

for $k = 1, ..., n-1$. However, note that the number of equations is higher by number 2 which makes us to choose $m_0 = m_n = 0$. Finally, we use these values and then equations

$$y_k = \frac{m_k}{6}h_k^2 + p_k h_k, \quad y_{k+1} = \frac{m_{k+1}}{6}h_k^2 + q_k h_k$$

to find $p_k, q_k's$. We have the following algorithm

Algorithm 7 Algorithm for cubic spline.

(1) Set $m_0 = m_n = 0$

(2) Solve m_k

$$m_k h_k + 2m_{k+1}(h_k + h_{k+1}) + m_{k+2}h_{k+1} = 6(d_{k+2} - d_{k+1}),$$

(3) Solve p_k and q_k,

$$p_k = \frac{f(x_k)}{h_k} - \frac{m_k}{6}h_k$$

and

$$q_k = \frac{f(x_{k+1})}{h_k} - \frac{m_{k+1}}{6}h_k$$

(***)

$$h_{k-1}m_{k-1} + 2(h_{k-1} + h_k)m_k + h_k m_{k+1} = 6(d_k - d_{k-1}),$$

Let's write (***) in the matrix form. (3阶)

One can solve the cubic spline problem for large dimensions with the Jacobi or Gauss-Seidel method. For that note that we can write the problem in the following form:

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ h_0 & 2(h_0+h_1) & h_1 & \dots & 0 \\ 0 & h_1 & 2(h_1+h_2) & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}}_{A} \begin{bmatrix} m_0 \\ m_1 \\ m_2 \\ \dots \\ m_n \end{bmatrix} = \begin{bmatrix} 0 \\ 6(d_2 - d_1) \\ \dots \\ 6(d_n - d_{n-1}) \\ 0 \end{bmatrix}$$

Linear Algebra.
Strictly diagonally domaint ⇒ there is a unique solution

The first matrix satisfies the condition for Jacobi method ($|a_{ii}| \geq \sum_{j \neq i} a_{ij}$) as $2|h_k + h_{k+1}| > |h_k + h_{k+1}|$. So one can use Jacobi method to find m_i's.

Example 20. Let us consider that from a function $\boxed{f(x) = \frac{1}{1+x^2},}$ we only have access to three values at $x = -1, 0, 1$ of the function. We want to use cubic interpolation using piece-wise polynomials of degree 3 in the following form

$$S_k(x) = \frac{m_k}{6h_k}(x_{k+1} - x)^3 + \frac{m_{k+1}}{6h_k}(x - x_k)^3 + p_k(x_{k+1} - x) + q_k(x - x_k),$$
$$\text{on } [x_k, x_{k+1}].$$

We have to find m_1, q_k and p_k, $k = 0, 1,$ by assuming $\boxed{m_0 = m_2 = 0.}$ Observe that $\boxed{h_0 = h_1 = 1.}$ So first see that

By Boundary assumption.

distance between x_0, x_1 x_2, x_3

$$d_1 = \frac{f(x_1) - f(x_0)}{h_0} = \frac{1 - \frac{1}{2}}{1} = \frac{1}{2},$$

$$d_2 = \frac{f(x_2) - f(x_1)}{h_1} = \frac{\frac{1}{2} - 1}{1} = -\frac{1}{2}.$$

Now given $m_0 = m_2 = 0$, $h_0 = h_1 = 1$, and the values of $d_1$and d_2, in the algorithm above we get

$h_0 = h_1 = 1.$

$$\cancel{m_0 h_0} + 2m_1 \underbrace{(h_0 + h_1)}_{2.} + \cancel{m_2 h_1} = 6\left(d_2 - d_1\right),$$

$$\Rightarrow 4m_1 = -6 \Rightarrow m_1 = -\frac{3}{2}.$$

Now we solve for p_k, q_k

$$p_0 = \frac{f(x_0)}{h_0} - \frac{m_0}{6}h_0 = \frac{1}{2},$$

$$p_1 = \frac{f(x_1)}{h_1} - \frac{m_1}{6}h_1 = \frac{5}{4},$$

$$q_0 = \frac{f(x_1)}{h_0} - \frac{m_1}{6}h_0 = \frac{5}{4},$$

$$q_1 = \frac{f(x_2)}{h_1} - \frac{m_2}{6}h_1 = \frac{1}{2}.$$

plugg in in.

So by replacing the values obtained in the cubic spline formula we get

$$S_0(x) = \frac{m_0}{6h_0}(x_1 - x)^3 + \frac{m_1}{6h_0}(x - x_0)^3 + p_0(x_1 - x) + q_0(x - x_0), \text{ on } [x_0, x_1],$$

$$\Rightarrow S_0(x) = -\frac{1}{4}(x+1)^3 - \frac{1}{2}x + \frac{5}{4}(x+1), \text{ on } [-1, 0],$$

and $= f(0) = 1 \Rightarrow S(0) = -\frac{1}{4} + \frac{5}{4}$

$$S_1(x) = \frac{m_1}{6h_1}(x_2 - x)^3 + \frac{m_2}{6h_1}(x - x_1)^3 + p_1(x_2 - x) + q_1(x - x_1)$$

$$\Rightarrow S_1(x) = -\frac{1}{4}(1-x)^3 + \frac{5}{4}(1-x) + \frac{1}{2}x, \text{ on } [0, 1].$$

$= f(1) = \frac{1}{2}$

Remark: [Error Analysis]

Let $f \in C^4[a,b]$, with $M_4 = \sup |f^{(4)}(x)|$. Let S be the unique cubic spline interpolating polynomial at data points,
$|f(x) - S(x)| \le \frac{5M}{384} \max_{1\le k\le n-1} (x_k - x_{k-1})^4$

Non composite method $\left|\int_a^b f(x)dx\right|$.

composite $\left| f \approx P \Rightarrow \int f dx \approx \int P dx \right|$

1.5. Numerical Integral

1.5.1. Introduction. Finding the integral of a function is always a very important topic. In finance in particular, there are many times that we need to find the integral of a function with no closed form solution. For instance, finding the cumulative distribution function of a standard normal distribution plays a crucial role in pricing put and call options. That means we always need to estimate the value of the following integral:

$$\begin{cases} \Phi(x) = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{s^2}{2}} ds, & \text{for } \ x \geq 0. \\ \Phi(x) = 1 - \Phi(-x) & \text{for } x \leq 0 \end{cases}$$

But there is no close form for this integral, so we need to find an estimation of that. In the following we introduce different methods for numerical integration, that is based on the idea of polynomial interpolation. More precisely, we estimate the integral when the real function is interpolated by a polynomial, for which we know its integral. So we introduce two type of methods:

(1) Non-composite methods

(2) Composite methods

1.5.2. Non-composite methods. $n \geq 1$. Let us consider $f : [a,b] \to R$. Let us interpolate the function f given values at point $x_0 = a$ and $x_k = x_0 + kh$, $k = 0, 1, ..., n$, where $h = \frac{b-a}{n}$. For simplicity, let us denote $f_k = f(x_k) = f(x_0 + kh), k = 0, 1, ..., n$. Then, using the Lagrangian polynomials we can use the polynomial interpolation to estimate the function f: Let's $P = P_n$ denotes the nth Lagarlange Interolating polynomial of f at these data point.

$$f(x) \approx P_n(x) = \sum_{k=0}^{n} f_k L_k(x)$$

Therefore, the integral $\int_a^b f(x)dx = \int_{x_0}^{x_n} f(x)dx$ can be found as

$$\begin{aligned} \int_{x_0}^{x_n} f(x)dx &\approx \int_{x_0}^{x_n} P_n(x)dx = \int_{x_0}^{x_n} \sum_{k=0}^{n} f_k L_k(x)dx \\ &= \sum_{k=0}^{n} f_k \int_{x_0}^{x_n} L_k(x)dx = \sum_{k=0}^{n} f_k w_k, \end{aligned}$$

$\begin{cases} x_n = b \\ x_0 = a \end{cases}$

w_k (完全 Independent to f_k) weighted sums.

where $w_k = \int_{x_0}^{x_n} L_k(x) = \int_a^b L_k(x)$. So the integral of a function is approximated by a weighted sum of the values of the function f at $a =$

$x_0, \ldots, x_n = b$. It is very important to note that the weights $w_k, k = 0, \ldots, n$ does not depend to f. It is also interesting that w_k's also are only a function of h.

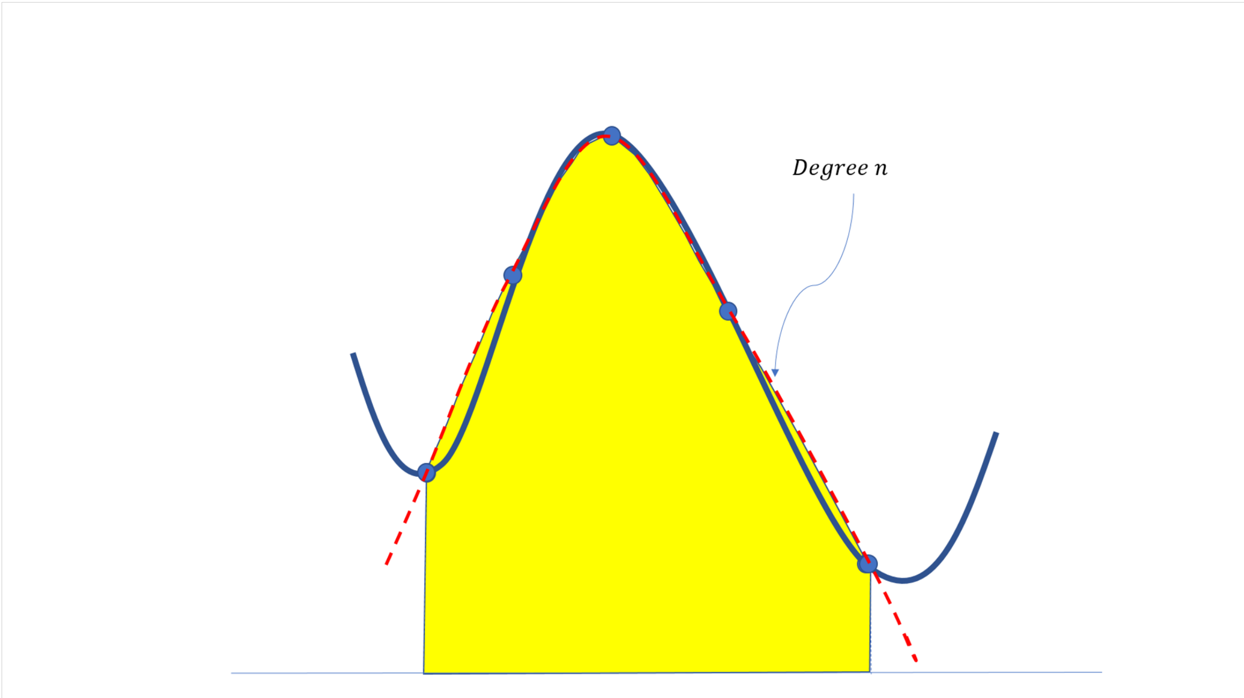


FIGURE 1.5.1. Integral with polynomial interpolation.

In this lecture we only consider $n = 1, 2$. By using the same notation as above we have,

$a = x_0 < x_1 = b$, $h = b - a$, $W_0 = \int_a^b \left(\frac{x-a}{a-b}\right) dx = \frac{h}{2}$, $W_1 = \int_a^b L_1(x) dx = \int_a^b \left(\frac{x-a}{b-a}\right) dx$

(1) if $n = 1 \Longrightarrow w_0 = w_1 = \frac{h}{2}$.

$\int_a^b f(x) dx \approx \frac{h}{2}(f(x_0) + f(x_1)) = \frac{h}{2}(f(a) + f(b))$

(2) if $n = 2 \Longrightarrow w_0 = w_2 = \frac{h}{3}, w_1 = \frac{4}{3}h$.

$a < x_0 < x_1 = \frac{a+b}{2} < x_2 = b$, $W_0 = \int_a^b L_0(x) dx$
$= \int_a^b \frac{(x-x_1)(x-b)}{(a-x_1)(a-b)} dx$
$= \frac{h}{3}$.

(3) if $n = 3 \Longrightarrow w_0 = w_3 = \dfrac{3}{8}h, \ w_1 = w_2 = \dfrac{9}{8}h$.

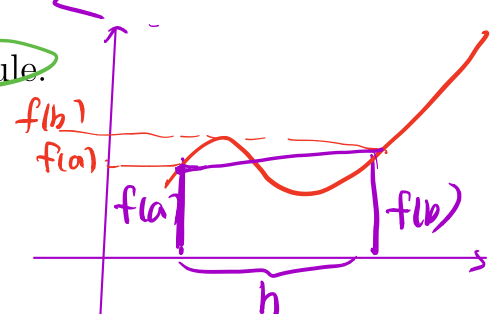
(4) if $n = 4 \Longrightarrow w_0 = w_4 = \dfrac{14}{45}h, \ w_1 = w_3 = \dfrac{64}{45}h, \ w_2 = \dfrac{24}{35}h$.

Therefore, using these weights we can introduce the following rules:

(1) $\int_{x_0}^{x_1} f(x)dx \approx \frac{h}{2}(f_0 + f_1)$, trapezoidal rule.

(2) $\int_{x_0}^{x_2} f(x)dx \approx \frac{h}{3}(f_0 + 4f_1 + f_2)$, Simpson rule.

$\int_a^b f(x) dx \approx$
$\frac{h}{3}(f(x_0) + 4f(x_1) + f(x_2))$
$= \frac{h}{3}(f(a) + 4f(a+h) + f(b))$



(3) $\int_{x_0}^{x_3} f(x)dx \approx \dfrac{3h}{8}(f_0 + 3f_1 + 3f_2 + f_3)$, Simpson 3/8 Rule.

(4) $\int_{x_0}^{x_4} f(x)dx \approx \dfrac{2h}{45}(7f_0 + 32f_1 + 12f_2 + 32f_3 + 7f_4)$, Boole's Rule.

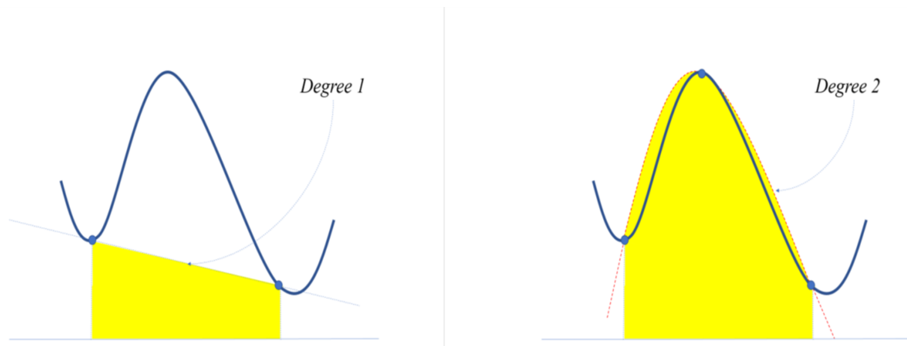


FIGURE 1.5.2. Integral with polynomial interpolation, trapezoidal and Simpson's rule.

Let us show how we can find these values. First, let us do it for $n = 1$. Without loss of generality, just by a transformation, assume that $a = 0$. For that we need to find first the Lagrangian functions

$$L_0(x) = -\frac{x-b}{b}, L_1(x) = \frac{x}{b}.$$

Therefore, given that $h = b - a = b$, we have

$$\begin{aligned} w_0 &= \int_a^b L_0(x)dx = \int_0^b \frac{b-x}{h}dx \\ &= \frac{1}{h}\left(bx - \frac{x^2}{2}\right)\Bigg|_{x=0}^{x=b} = \frac{1}{h}\left(b^2 - \frac{b^2}{2}\right) = \frac{h}{2} \end{aligned}$$

and

$$\begin{aligned} w_1 &= \int_a^b L_0(x)dx = \int_0^b \frac{x}{h}dx \\ &= \frac{1}{h}\frac{x^2}{2}\Bigg|_{x=0}^{x=b} = \frac{1}{h}\frac{b^2}{2} = \frac{h}{2}. \end{aligned}$$

$h = \frac{b-a}{n}$.

Now let us do it for $n = 2$. Again without loss of generality, and by transformation, assume $x_0 = -b, x_1 = 0, x_2 = b$, where $h = b$.

$$L_0(x) = \frac{x^2 - bx}{2b^2}, L_1(x) = \frac{x^2 - b^2}{-b^2}, L_2(x) = \frac{x^2 + bx}{2b^2}.$$

So we have

$$\begin{aligned} w_0 &= \int_{-b}^{b} L_0(x)dx = \int_{-b}^{b} \frac{x^2 - bx}{2b^2} dx \\ &= \frac{1}{2b^2} \left(\frac{x^3}{3} - \frac{bx^2}{2} \right) \Bigg|_{x=-b}^{x=b} = \frac{1}{2b^2} \left(\frac{2b^3}{3} \right) = \frac{h}{3} \end{aligned}$$

and

$$\begin{aligned} w_1 &= \int_{-b}^{b} L_1(x)dx = \int_{-b}^{b} \frac{x^2 - b^2}{-b^2} dx \\ &= \frac{1}{-b^2} \left(\frac{x^3}{3} - b^2 x \right) \Bigg|_{x=-b}^{x=b} = \frac{1}{-b^2} \left(\frac{2b^3}{3} - 2b^3 \right) = \frac{4h}{3} \end{aligned}$$

and

$$\begin{aligned} w_2 &= \int_{-b}^{b} L_2(x)dx = \int_{-b}^{b} \frac{x^2 + bx}{2b^2} dx \\ &= \frac{1}{2b^2} \left(\frac{x^3}{3} + \frac{bx^2}{2} \right) \Bigg|_{x=-b}^{x=b} = \frac{1}{2b^2} \left(\frac{2b^3}{3} \right) = \frac{h}{3}. \end{aligned}$$

$E(f) = \frac{1}{(n+1)!} \left| \int_a^b \prod_{i=0}^{n} (x - x_i) f^{(n+1)}(\xi(x)) dx \right|$

$\leq \frac{M_{n+1}}{(n+1)!} \left| \int_a^b \prod_{i=0}^{n} (x - x_i) dx \right|$, $M_n = \sup |f^{(n)}(x)|$, $x \in [a,b]$

Error terms. The error terms for a smooth enough function is given as (without proof)

$n = 1$, $|E(f)| \leq \frac{M_2}{2} h^3$

- Trapezoidal rule $\left| \int_{x_0}^{x_1} f(x)dx - \frac{h}{2}(f_0 + f_1) \right| \leq D_1 M_2 h^3$

If we further assume f is 4 times continously differentiable, i.e., $f \in C^4[a,b]$, then

- Simpson's rule $\left| \int_{x_0}^{x_2} f(x)dx - \frac{h}{3}(f_0 + 4f_1 + f_2) \right| \leq \underset{90}{D_2} M_4 h^5$

$3 \times 5 \times 6$

for some positive numbers $D_1$and D_2 where $M_k = \sup \left| f^{(k)}(x) \right|$.

For the Simpson 3/8 Rule and Boole's Rule we also have

- Simpson 3/8 Rule: $\left| \int_{x_0}^{x_3} f(x)dx - \frac{3h}{8}(f_0 + 3f_1 + 3f_2 + f_3) \right| \leq D_3 M_4 h^3$

Remark: when f is a polynomial of degree, then Simpson rule will provide an exact approximation

Example: $\int_0^2 \sin x \, dx$; (a) by Trapezoidal (b) Simpson's Rule.

(a) $\int_0^2 \sin x\,dx \approx (f(0)+f(2)) \approx 0.909$

(b) $\int_0^2 \sin x\,dx \approx \frac{1}{3}(f(0)+4f(1)+f(2)) \approx 1.425$

Exact solution is 1.416

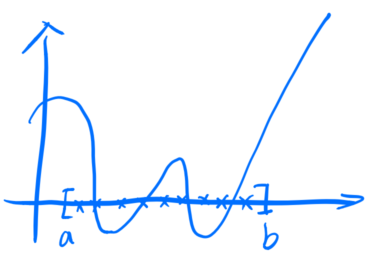
- Boole's Rule: $|\int_{x_0}^{x_4} f(x)dx - \dfrac{2h}{45}(7f_0+32f_1+12f_2+32f_3+7f_4)| \leq D_4M_6h^7$

1.5.3. Composite method. In order to find a more accurate integral we do not need to find a higher degree polynomial. We need just to choose a lower degree (like $n = 1, 2, 3$) and try to find integral of smaller intervals, separately and add them up. $\int_a^b f(x)\,dx = \sum_{k=1}^{n} \int_{x_{k-1}}^{x_k} f(x)\,dx \approx \sum_{k=1}^{n} \frac{h}{2}(f(x_{k-1}) + f(x_k))$ trapezidal rule.

Composite Trapezoidal rule. Let $a = x_0, ..., x_n = b$ be a partition for $[a, b]$, then we have

$$\begin{aligned}
\int_{x_0}^{x_n} f(x)dx &= \int_{x_0}^{x_1} f(x)dx + \int_{x_1}^{x_2} f(x)dx + ... + \int_{x_{n-2}}^{x_{n-2}} f(x)dx + \int_{x_{n-1}}^{x_n} f(x)dx \\
&\approx \frac{h}{2}(f_0+f_1) + \frac{h}{2}(f_1+f_2) + ... + \frac{h}{2}(f_{n-2}+f_{n-1}) + \frac{h}{2}(f_{n-1}+f_n) \\
&= \frac{h}{2}(f_0 + \underline{2f_1} + \underline{2f_2} + ...\underline{2f_{n-1}} + f_n).
\end{aligned}$$

composite trapezoidal Rule.



$\int_a^c f(x)dx = \int_a^b f(x)\,dx + \int_c^b f(x)\,dx.$

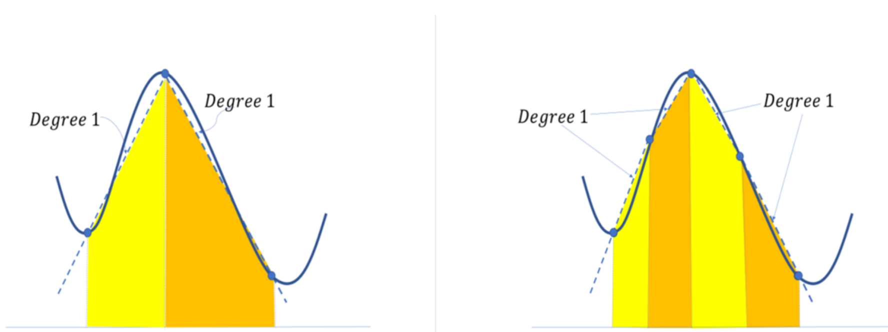


FIGURE 1.5.3. Integral with composite trapezoidal method.

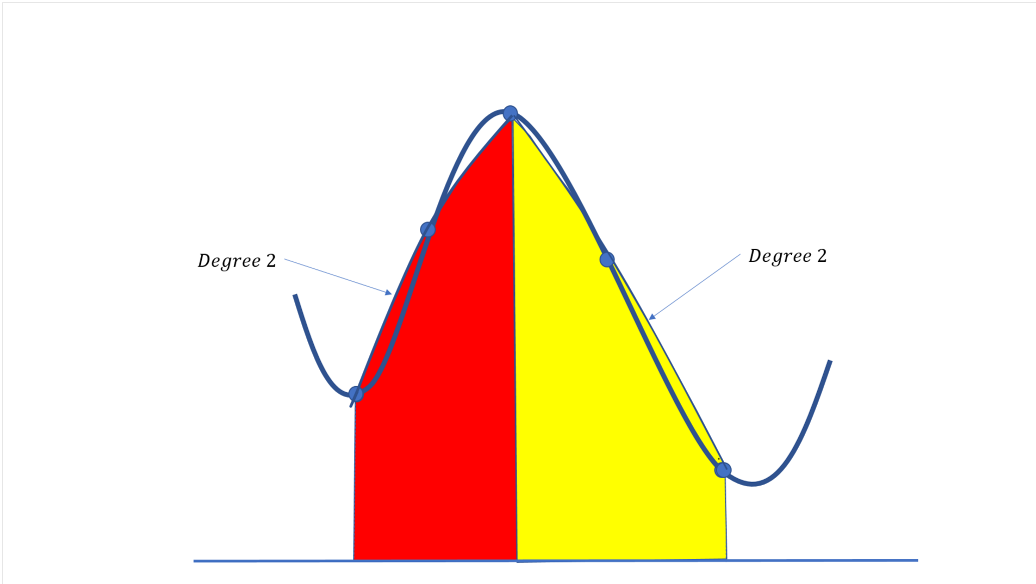


FIGURE 1.5.4. Integral with composite Simpson's rule.

EXAMPLE 21. Using the composite trapezoidal rule we find the value of $\Phi(1) = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \int_0^1 \mathrm{e}^{-\frac{s^2}{2}} ds$ for different partitions $n = 4, 8, 12$ and 16 and get the following results. To reach accuracy of 0.0000001, comparing to the real value we need to iterate 450 times.

$n = 4$	0.840081845
$n = 8$	0.841029516
$n = 12$	0.841204684
$n = 16$	0.841265969

FIGURE 1.5.5. Values of the integral using composite trapezoidal rule for different iterations.

Composite Simpson's rule. Note that for this rule we have to consider even numbers on the node since each piece of integral need three consecutive points. So consider $a = x_0, ..., x_{2m=n} = b$. We find integral on each $[x_{2k}, x_{2k+2}]$

Let's mimic the same idea:

a $b = x_{2m=n}$.

$$\int_a^b f(x)dx = \sum_{k=0}^{m-1} \int_{x_{2k}}^{x_{2k+2}} f(x)\,dx$$

$$\approx \sum_{k=0}^{m-1} \frac{h}{3}\left(f(x_{2k}) + 4f(x_{2k+1}) + f(x_{2k+2})\right)$$

$$= \frac{h}{3}\left(f(x_0) + 2\left(f(x_2) + f(x_4) + \cdots f(x_{2m-2})\right) + 4\left(f(x_1) + f(x_3) + \cdots + f(x_{2m-1})\right) + f(x_{2n})\right)$$

$$\int_{x_0}^{x_n} f(x)dx = \int_{x_0}^{x_{2m}} f(x)dx = \int_{x_0}^{x_2} f(x)dx + \int_{x_2}^{x_4} f(x)dx + \ldots + \int_{x_{2m-2}}^{x_{2m}} f(x)dx$$
$$\approx \frac{h}{3}(f_0 + 4f_1 + f_2) + \ldots + \frac{h}{3}(f_{2m-2} + 4f_{2m-1} + f_{2m})$$
$$= \frac{h}{3}(f_0 + 4(f_1 + f_3 + \ldots + f_{2m-1}) + 2(f_2 + \ldots + f_{2m-2}) + f_{2m})$$

EXAMPLE 22. Using the composite Simpson's rule we find the value of $\Phi(1) = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \int_0^1 e^{-\frac{s^2}{2}} ds$ for different partitions $n = 4, 8, 12$ and 16 and get the following results. To reach accuracy of 0.0000001, comparing to the real value we need to iterate 16 times.

$n = 4$	0.841355488
$n = 8$	0.841345406
$n = 12$	0.841344876
$n = 16$	0.841344876

FIGURE 1.5.6. Values of the integral using composite Simpson's rule for different iterations.

Error term for composite method. Let us find error term for composite trapezoidal and Simpson's rules. For the trapezoidal rule we need to sum up the errors of each sub-interval:

(a).

$$|E(f)| = \left| \int_a^b f(x)dx - \underbrace{\frac{h}{2}(f_0 + 2f_1 + 2f_2 + \ldots 2f_{n-1} + f_n)}_{\sum_{k=1}^{n} \frac{h}{2}(f(x_{k-1}) + f(x_k))} \right| \le \sum_{k=1}^{n} \left| \int_{x_{k-1}}^{x_k} f(x)dx - \frac{h}{2}(f(x_{k-1}) + f(x_k)) \right|$$
$$\le \sum_{i=1}^{n} \underset{\downarrow \frac{1}{12}}{D_1} M_2 h^3 = D_1 M_2 n h^3 = D_1 M_2 \frac{(b-a)^3}{n^2}.$$
$$= \frac{M_2}{12}(b-a)^3 n^{-2}$$

The same is true for Simpson's rule for $n = 2m$ point,

$$|E_{cs}(f)| = \left| \int_a^b f(x)dx - \frac{h}{3}(f_0 + 4(f_1 + f_3 + \ldots + f_{2m-1}) + 2(f_2 + \ldots + f_{2m-2}) + f_{2m}) \right|$$
$$\le \sum_{i=1}^{n} \underset{\downarrow \frac{1}{90}}{D_2} M_4 h^5 = D_2 M_4 \underset{\frac{n}{2}}{n} h^5 = D_2 M_4 \frac{(b-a)^5}{n^4}.$$
$$= \frac{M_4}{180}(b-a)^5 n^{-4}.$$

Example: Determine the minimum number n that will ensure an error of less than 2×10^{-5} when approximately $\int_0^{\pi} \sin x\,dx$

(a) Composite Trapezoidal (b) Composite Simpson Rule.

solution (a)

$\left|\frac{\pi}{12}f''(x)\right| n^{-2} \leq 2\times10^{-5}$. $M_2 = \sup|f''(x)| \leq 1,\ x\in[0,\pi]$

$\frac{\pi^3}{12}n^{-2} \leq 2\times10^{-5} \Rightarrow n \geq \left(\frac{\pi^3}{12\times2\times10^{-5}}\right)^{\frac{1}{2}} \approx 359.44,\ n \geq 360$

(b) $\frac{\pi^5}{180}n^{-4} < 2\times10^{-5} \Rightarrow n > \left(\frac{\pi^5}{180\times2\times10^{-5}}\right)^{\frac{1}{4}} \approx 19.09 \Rightarrow n \geq 18$

1.5.4. Runge phenomena. Range phenomena happens when approximation with higher degree polynomials does not help for accuracy. Reminding that when we approximate f by a polynomial P_n thee error term at each point is,

$$E(x) = \frac{(x-x_0)(x-x_1)...(x-x_n)f^{n+1}(c)}{(n+1)!}$$

But when $n \to \infty$ it might be thee case that $f^{n+1}(c)$ grows by n and the fraction $E(x)$ never gets smaller and smaller. For example let us look at function,

$$f(x) = \frac{1}{1+25x^2}$$

and try to approximate at $[-1,1]$. Then for equal spaces intervals we will assume $x_i = -1 + (i-1)\frac{2}{n}$, $(i = 1, 2, ..., n+1)$,

$$f(x) - P_n(x) = \frac{f^{n+1}(\zeta)}{(n+1)!}\Pi_{i=0}^{n}(x - x_i)$$

But let us see,

$|f'(1)| = 0.074$

$|f''(1)| = 0.2105$

$|f'''(1)| = ...$ Larger and larger.

1.6. Exercise

EXERCISE 23. Suppose $F(x) = 5x^2$. By using root-finding numerical method we want to find the numerical inverse of F at $F(x) = 1, 2, 3, 4, 5$ when $x \in [0, 1]$. In order to do so, for any $y = 1, 2, 3, 4, 5$ we solve the equation $F(x) = y$ in terms of x. Using the bisection method the following script is an incomplete program to find x. Please complete it and print out the Matlab results.

```
clear all

for i=1:5                                          % for loop to iterate all y's
    tol=0.00000001; a=0; b=1; delta=1; % declare all the parameters
while (delta>tol)                % start iterate for bisection method
    % answer: c=(a+b)/2;
    f1=5*(a^2)-i;
    % answer: f2=5*(c^2)-i;
    f3=5*(b^2)-i;
    if (suggested place)
        b=c;
    else
        (suggested place);
    end
    % answer: delta = abs(5*(c^2)-i);

end

end
```

EXERCISE 24. Let us assume we want to find $\sqrt{2}$. For that we have to find the solution of $x^2 = 2$. Use the Newton method to find a numerical method to solve this problem and report the first iteration when the initial point is 1.4.

EXERCISE 25. Let $L > 1$. We want to find the minimum of the function $f(x) = (x - a)^L$ by using the Newton method.

(1) Denoting the Newton iterative sequence by $\{x_k\}_{k=0,1,2,\ldots}$, find x_{k+1} in terms of x_k.

(2) If x_0 is chosen so that x_k is convergent, find the limit of the sequence x_k, and show how you have found this limit.

(3) Show by induction that

$$x_k = \left(\frac{L-1}{L}\right)^k (x_0 - a) + a.$$

Using this relation, find the root of f.

(4) Using this relation, for $L = 4$, $a = 1$, $x_0 = \frac{1}{4}$, find the minimal point x. Using part 2, find the smallest i such that the error term $|x_i - x|$ is smaller than or equal to $\delta = 0.0001$.

EXERCISE 26. The following questions are about Nelder-Mead. Please give details for the first iteration and the corresponding two simplexes.

a) For function $f(x, y) = x^2 - 4x + y^2 - y - xy$, starting from $\overrightarrow{x_1} = (2.4, 1.6)$, $\overrightarrow{x_2} = (2.4, 2.4)$ and $\overrightarrow{x_3} = (3.6, 1.6)$.

b) For function $f(x, y) = |x| + |y|$, $(x, y) \in \mathbb{R}^2$, starting from $\overrightarrow{x_1} = (0, 0)$, $\overrightarrow{x_2} = (1, 0)$ and $\overrightarrow{x_3} = (0, 1)$.

c) For function $f(x, y) = \sqrt{1 - x} + |y - 1|$ starting from $\overrightarrow{x_1} = (3/4, 1)$, $\overrightarrow{x_2} = (1, 2/3)$ and $\overrightarrow{x_3} = (1, 0)$.

d) For function $f(x, y) = x^2 + y^2$, starting from $\overrightarrow{x_1} = (1, 0)$, $\overrightarrow{x_2} = (0, 1)$ and $\overrightarrow{x_3} = (2/5, 2/5)$.

e) For function $f(x, y) = x^2 + y^2$ starting $\overrightarrow{x_1} = (0, 1)$, $\overrightarrow{x_2} = (1, 0)$ and $\overrightarrow{x_3} = (1, 1)$.

EXERCISE 27. For the same function $f(x, y) = x^2 - xy + y^2 - 4x - y$,

$$\nabla f(x, y) = (2x - y - 4, 2y - x - 1)$$

please use Gradient Method (Steepest Descent Method) to find its minimum point manually with initial point $x_0 = (3, 3)$ for one iteration.

Exercise 28. Consider the following function

$$f(x,y) = e^{(x-4)^2} + (y-2)^2 .$$

We want to find the minimum of this function using the gradient method (or steepest decent). Starting from $\overrightarrow{x_0} = (3,1)$. Using $\gamma = 0.1, 0.2, 0.3$ find the second value of the function f at the second iteration and say what are the points they take this values.

Exercise 29. Consider the following system of linear equations

$$\begin{cases} 3x_1 + x_4 & = 1 \\ 3x_2 & = 0 \\ 3x_3 & = 0 \\ x_1 + 3x_4 & = 1 \end{cases} \tag{1.6.1}$$

a) Write this system of equations in matrix form

$$A\overrightarrow{x}^T = b.$$

b) Decompose A to lower triangular, upper triangular and diagonal matrices L, U and D, respectively.

c) Using the Jacobi method, what is the function that whose fixed point is the answer to system (1).

d)If we denote the sequence produced by Jacobi method by

$$\left\{\overrightarrow{x}_k = \left(x_1^k, x_2^k, x_3^k, x_4^k\right)\right\}_{k=0,1,2,\ldots}$$

find $\overrightarrow{x}_k$ iteratively (i.e., $\overrightarrow{x}_{k+1}$ in terms of $\overrightarrow{x}_k$).

e)Using part d), show that

$$\begin{aligned} x_1^{k+1} &= \tfrac{2}{9} + \frac{x_1^{k-1}}{9} \\ x_4^{k+1} &= \tfrac{2}{9} + \frac{x_4^{k-1}}{9} \end{aligned}$$

Using these relations and part d), find the answer to system (1).

EXERCISE 30. Consider the following system of linear equations

$$\left\{\begin{array}{ll} 2x_1 + x_3 & = 2 \\ x_2 & = 0 \\ x_1 + 2x_3 & = 1 \end{array}\right. \tag{1.6.2}$$

Write this system of equations in matrix form

$$A\overrightarrow{x}_k^T = b.$$

(1) Decompose A into lower triangular, upper triangular and diagonal matrices L, U and D, respectively.

(2) Using the Gauss-Seidel method, what is the function whose fixed point is the answer to system (1).

(3) If we denote the sequence produced by Jacobi method by

$$\left\{\overrightarrow{x}_k = \left(x_1^k, x_2^k, x_3^k\right)\right\}_{k=0,1,2,...}$$

find $\overrightarrow{x}_k$ iteratively (i.e., $\overrightarrow{x}_{k+1}$ in terms of $\overrightarrow{x}_k$), and find the solution to the system of equations (1) by sending $k \to \infty$. What is the smallest k so that we get an answer as close as $\delta > 0$, for a given positive number δ?

(4) Now using the Jacobi method, if we denote the sequence by

$$\left\{\overrightarrow{x}_k^* = \left(x_1^{*k}, x_2^{*k}, x_3^{*k}\right)\right\}_{k=0,1,2,...}$$

find $\overrightarrow{x}_k^*$ iteratively (i.e., $\overrightarrow{x}_k^*$ in terms of $\overrightarrow{x}_k^*$), and find the solution to the system of equations (1) by sending $k \to \infty$.

EXERCISE 31. Let us consider that from a function $y = f(x)$, we only have access to three values $y = 10, 20, 25$ of the function associated with $x = 10, 20, 30$, respectively. We want to find a curve which passes through these points. Answer the following questions.

a) Assume we want to pass a polynomial through all these points. What is the minimum degree of the polynomial? Using Legendre polynomials, find the polynomial which fits the data above.

b) Now use cubic spline to fit a curve to data.

Hint: Use a polynomial in the following form

$$S_k(x) = \frac{m_k}{6}(x_{k+1} - x)^3 + \frac{m_{k+1}}{6}(x - x_k)^3 + p_k(x_{k+1} - x) + q_k(x - x_k),$$

first match S_k'''s then S_k''s and then S_k's.

c) Use part a) and b) to compute the area below the curve, and compare two answers.

EXERCISE 32. In Example 20 use the cubic spline to estimate the integral $\int_{-1}^{1} f(x)dx$. Now we want to use polynomial interpolation instead. After finding Lagrangian polynomials L_0, L_1, L_2, find the degree two polynomial P_2 interpolating $(-1, f(-1))$, $(0, f(0))$ and $(1, f(1))$. Use P_2 to approximate $\int_{-1}^{1} f(x)dx$. By finding the true value of $\int_{-1}^{1} f(x)dx$, compare the error of the integral in the previous parts, and say which interpolation gives better approximation of the integral.

EXERCISE 33. Let $f : [1/1000, 8] \to \mathbb{R}$ be given by $f(x) = \sqrt[3]{x}$. For $j = 0, 1, 2, 3$ we have the following table:

j	x_j	$f(x_j)$
0	1/1000	1/10
1	1/8	1/2
2	1	1
3	8	2

a) Assume we want to pass a polynomial through all these points. What is the minimum degree of the polynomial? Write down the general form of Lagrange polynomials and use it to find the Lagrange polynomials for data above.
b) Write down the general form of a fitted polynomial using Lagrange polynomials, and use part a) to find the polynomial which fits the data above.
c) Use the polynomial you just created to approximate $\sqrt[3]{2}$ and compare

it with $\sqrt[3]{2} \approx 1.25992104$, how good is your approximation?
d) Use the error estimate given in the lecture to estimate the error here. What is the difference between the estimated error and the actual error?

Exercise 34. a) Let us assume that $g : \mathbb{R} \to \mathbb{R}$ is an absolutely increasing and smooth function. We need to find the inverse of the function g at point y i.e. $g^{-1}(y)$. For that, we will have to solve $f(x) = g(x) - y = 0$ by using Secant method. In order to find x, we write the following program. But this program has three gaps and is incomplete (at three different places). Please complete the program in an appropriate way. (**Hint:** The places that you have to fill the gaps in the program are suggested.

```
tol=.00001;
p0=x;
p1=y;
Suggested Place 1
while (delta>tol)
    p2=p0-((g(p0)-y)*(p1-p0)/(g(p1)-g(p0));
    delta=abs(g(p2)-y);
    Suggested Place 2
    Suggested Place 3
end
```

b) Now consider that we want to find $g^{-1}(y)$ by Newton method. Modify the program above and write down a program for Newton method (denote the derivative of g with g').

Exercise 35. a) We want to approximate $\log(l)$, for a given value $l > 1$, by writing a simple program which uses Newton's method to find the root of $f(x) = e^x - y$. We write the following program, but this program has two gaps and is incomplete (at two different places). Please complete the program in appropriate way. (**Hint:** The places that you have to fill the gaps in the program are suggested).

```
tol=.00001;
p0=1;
Suggested Place 1
while (delta>tol)
    p1=p0-(exp(p0)-l)/exp(p0);
    delta=abs(p1-p0);
```

```
    Suggested Place 2
end
```

b) Now consider that we want to find $\log(l)$, for $l > 1$, by using Secant method. Modify the program above and write down a program for Secant method.

c) Suppose $F : [x_0, x_1] \to [x_0, x_1]$ is an absolutely increasing function such that $F(x_0) = x_0$ and $F(x_1) = x_1$. By using root-finding numerical method we want to find the numerical inverse of F. In order to do so, for any $y \in [x_0, x_1]$ we solve the equation $F(x) = y$ in terms of x. We use the bisection method to solve this equation. The following program is an incomplete program to find x. Please complete it. (**Hint:** The places that you have to fill the gaps in the program are suggested.

```
tol=.00001;
a=x0;
b=x1;
delta=1;
while (delta>tol)
    Suggested place 2
    f1=Suggested place 1
    f2=F(c)-y;
    f3=Suggested place 1
    if f1*f2<=0
        b=c;
    else
        a=c;
    end
    Suggested place 3
end
```

Def: [log normal distribution]

we say that a r.v. Y follows a log normal distribution with parameters μ, σ^2, if

$$Y = e^X; \quad X \sim N(\mu, \sigma^2)$$

Exercise (at home) $\forall n \geq 1$; show that

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

$n = 1$

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

$$E(Y^2) = e^{2\mu + 2\sigma^2}.$$

$$Var(Y) = E(Y^2) - E(Y)^2.$$

$$= (e^{2\mu + \sigma^2})(e^{\sigma^2} - 1)$$

The message of this chapter, is to approximate price of popular options in (B.S.) market by using pricing in simple discrete market model

Def: [Normal (Gaussian) Distribution)

$$X \sim N(\mu, \sigma^2)$$

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

Black-Schole, Market Model.

① Non-Risky Asset (Bank account B)

$B_t = B_0 e^{rt}$ → Interest rate

② Risky Asset (stock S)

$dS_t = \mu S_t\,dt + \sigma S_t\,dW_t$ → Brownian Motion.

drift parameter

volatility parameter

$E(X)=\mu$, $Var(X)=\sigma^2$.

Remark: (a) when $\mu=0$, $\sigma^2=1$ $\Rightarrow$ X: standard ...

(b) If $X \sim N(\mu, \sigma^2)$; then $Y = \frac{X-\mu}{\sqrt{\sigma^2}} \sim N(0,1)$

(c) If $X \sim N(0,1)$, $Y = \sigma X + \mu \sim N(\mu, \sigma^2)$

(d) If $X_i \sim N(\mu_i, \sigma_i^2)$; $i=1,2$, and X_1, X_2 are indep.
then $X_1 + X_2 \sim (\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$

CHAPTER 2

Numerical methods in option pricing

2.1. Introduction and preliminaries

In this chapter the main target is to numerically find the price of derivatives, in particular, European and American option. More precisely we introduce methods to approximate the price of the derivatives. The approach that we take is to use multinational trees, in particular binomial and trinomial trees. One may have seen how to use binomial (or trinomial) trees to find the prices of derivatives, but it is important to note that we use them not as a model, but as a way to approximate a continuous time model. For that reason, we do not need to take care of the completeness of the markets with the multinational models, but about the approach that is used for approximation. The approach we use is known as the moment matching method, which equalizes the moments.

Let's start with reminding some basic concepts. A random variable is a variable which takes different values with different probabilities, due to different scenarios. For example if one has n scenarios numbered by 1 to n, whose probabilities are $p_1, ..., p_n$ respectively, then a random variable X is a variable that can take values $x_1, ..., x_n$ associated with these n scenarios and can be shown as follows:

$$X = \begin{cases} x_1 & \text{with} \quad p_1 \\ x_2 & \text{with} \quad p_2 \\ \vdots & \quad\vdots \qquad \vdots \\ x_n & \text{with} \quad p_n \end{cases}$$

Since the value of a random variable is uncertain, we rather want to have a number that can represent the value we expect to see from a random variable. That is why we introduce expectation.Mean or average (or expectation) is the average value of a r.v.,

$$E(X) = p_1x_1 + p_2x_2 + ... + p_nx_n.$$

However, the real value of a random variable, when it is observed, can be different from the expected value. That is why we introduce a variable that can measure how far the observed value can get from the expected value. Variance is to measure the distance from mean,

$$Var(X) = p_1(x_1 - E(X))^2 + p_2(x_2 - E(X))^2 + ... + p_n(x_n - E(X))^2.$$

Let us assume the price of an asset tomorrow is modeled by a simple random variable S_T which takes n different values. It is known that there is no arbitrage (risk free opportunity), then there is a probability measure $p_1, ..., p_n$ under which the discounted asset value is a martingale. It is not necessary in this lecture note to really know what that exactly means, but it is necessary to know if such a probability measure exists then the price of any contingent claim needs to be found under this probability. In this section, we simply assume that we are always endowed with such probability measure. However, the main concern is how to use a discrete model to approximate the price of contingent claims in continuous time models. For that reason we first need to review the discrete time models and then see how we can make a good approximation.

We always assume there are just two assets in the market, a risk asset and a risk free asset whose value is $e^{r\delta t}$ for one unit of cash.

The Binomial Market Model = (CRR)

2.2. Recombining binomial and trinomial trees

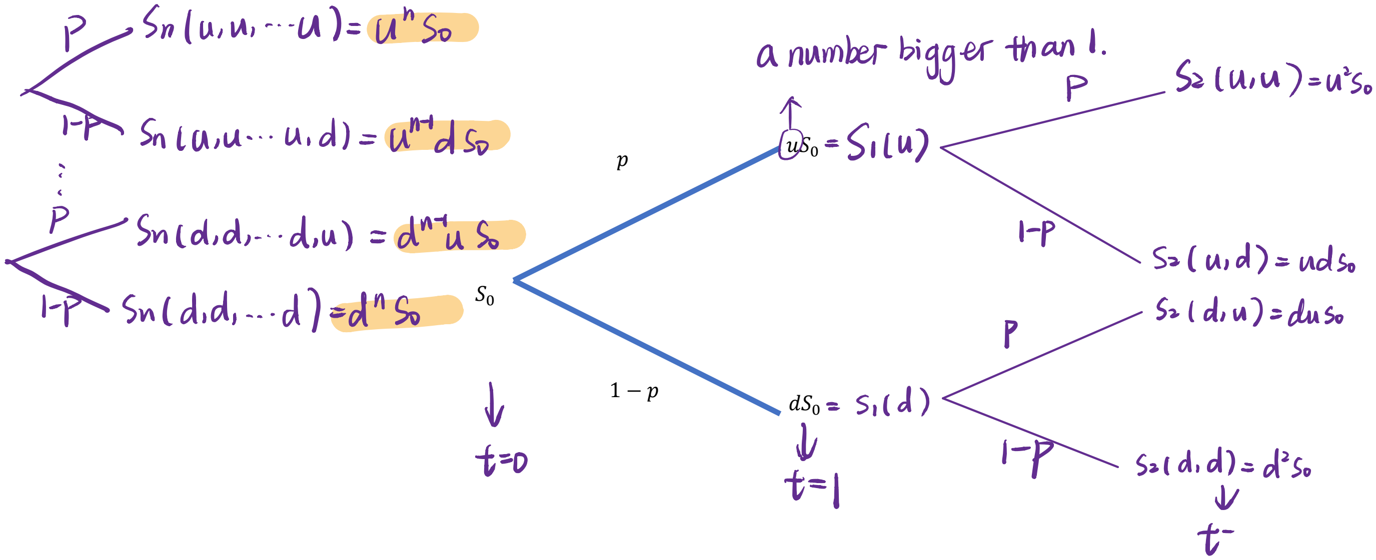
Let us assume today's price is S_0 and tomorrow price is either uS_0 with probability p or dS_0 with probability $1-p$ when $u > 1$ and $d < 1$ (with probability p it grows and $1-p$ it drops).

Discrete Model , $S_0 > 0$

$\downarrow$

$t=0$

Least Price at time n.



All the possible prices for the stock at time n are

$u^{n-k} d^k S_0$, $k=0,1,\cdots n$.

FIGURE 2.2.1. A simple binomial tree.

(n+1) different stock prices

Q: What is the probab. that price of stock at time n is given by $S_n = u^{n-k} d^k S_0$

However, in the real world we recombine this single so-called binomial trees and find a version that in n steps has $n+1$ out-put S_0u^n, $S_0u^{n-1}d$, ..., S_0ud^{n-1}, S_0d^n, with probabilities p^n, $\begin{pmatrix} n \\ 1 \end{pmatrix} p^{n-1}(1-p)$, ..., $\begin{pmatrix} n \\ n-1 \end{pmatrix} p(1-p)^{n-1}$, $(1-p)^n$.

$\binom{n}{i} p^{n-i} (1-p)^i$

Values	Probabilities
S_0u^n	p^n
$S_0u^{n-1}d$	$\begin{pmatrix} n \\ 1 \end{pmatrix} p^{n-1}(1-p)$
$\vdots$	$\vdots$
S_0ud^{n-1}	$\begin{pmatrix} n \\ n-1 \end{pmatrix} p(1-p)^{n-1}$
S_0d^n	$(1-p)^n$

TABLE 1. Probabilities of the last nodes of a recombining binomial tree.

For every $1 \le i \le n$, there are $(i+1)$ different stock price, and

$\forall\ 0 \le j \le i$

$S_i = u^{i-j} d^j S_0$

with probability $\binom{T}{j} p^{T-j} (1-p)^j$

Remark: $ud=1$

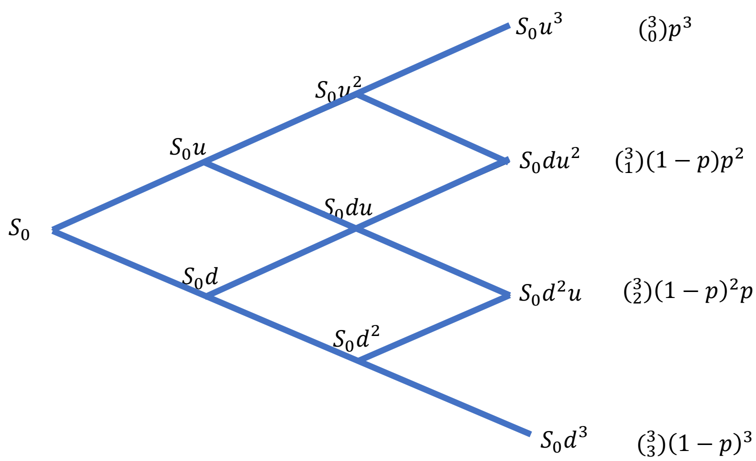
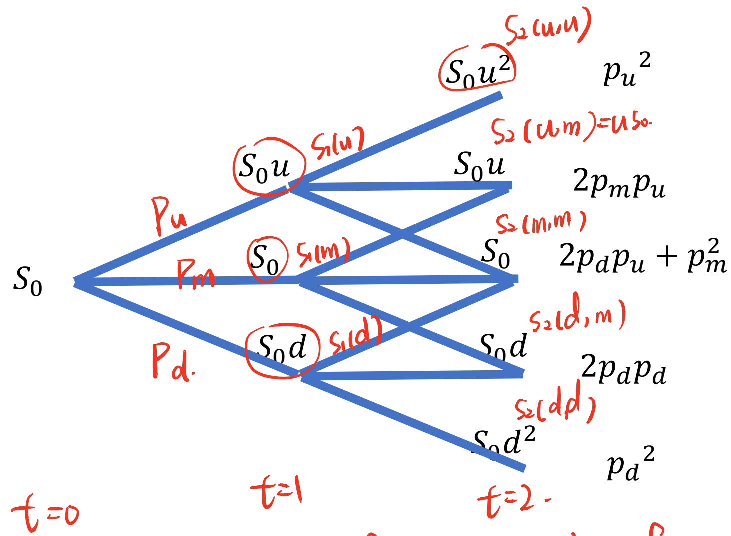


FIGURE 2.2.2. A recombining binomial tree.

$\Rightarrow S_T = u^{T-2j} S_0$ for $j=0,1,\cdots T$.

Like above, let us assume today's price is S_0 and tomorrow price is either uS_0 with probability p_u, dS_0 with probability p_d, or simply remains the same S_0, with probability p_m, when $u > 1$ and $d < 1$. In trinomial trees we always assume $u = 1/d$, that is why $ud = 1$ is always applied, like in the following figure (note that we usually need to find the probabilities in this case).



After n steps; we will have the following prices for stock:

$u^{N-k}S_0$; $0\le k<N-1$ (2N+1) different stock prices

$d^{N-k}S_0$; $0\le k<N-1$

S_0

FIGURE 2.2.3. A recombining trinomial tree.

Let us denote the value of an asset at period j and scenario i by $S_j(i)$, where we have n period i.e., $n\delta t = T$. Let's denote the value of the asset at scenario i at period n associated with time T by $S_T(i)$. For a recombining binomial tree at period i there is $i+1$ scenarios and for a recombining trinomial tree there is $2i+1$ scenarios.

THEOREM 36. *Assuming that* $u = 1/d$ *then*

(1) For a recombining binomial tree we have $S_j(i) = u^{j-2i+2}$ for $i = 1, ..., j+1$.

(2) For a recombining trinomial tree we have $S_j(i) = u^{j-i+1}$ for $i = 1, ..., 2j+1$.

Option Pricing in Binomial/Trinomial Market Model.

$D = f(S_T)$ for some given function

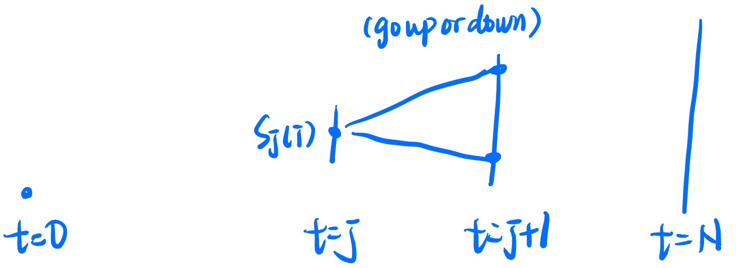
$D_j(i)$ is the value of Derivate D at time j at scenario i.

$0 \le j \le N$

$0 \le i \le j+1$, scenario i corresponds to the node that stock price take value with this notation

$j = N$; $D_N(i) = f(S_N(i)) = f(u^{N-i} d^i S_0)$, $0 \le i \le N$

$\pi_j^D(i)$ = Price of the option D at time j in scenario i.



$\pi_j^D(i)$

$$S_j(i) = (u^{j-i} d^i S_0) \times \begin{cases} u = u^{j+1-i} d^i S_0 = S_{j+1}(i) \\ d = u^{j-i} d^{i+1} S_0 = u^{(j+1)-(i+1)} d^{i+1} S_0 = S_{j+1}(i+1) \end{cases}$$

Therefore

$$\pi_j^{(D)}(i) = e^{-r\delta t}\left(p\,\pi_{j+1}^D(i) + (1-p)\,\pi_{j+1}^D(i+1)\right)$$

counting the same; one can obtain that

Example: [European call / put options]

$$\pi_j^{EC}(i) = e^{-r\delta t}\left(p\,\pi_{j+1}^{Ec}(i) + (1-p)\,\pi_{j+1}^D(i+1)\right)$$

when $j = N-1$,

$$\pi_j^{EC}(i) = e^{-r\delta t}\left(p(S_N(i)-k)_+ + (1-p)(S_N(i+1)-k)_+\right)$$

Also,

$$\pi_0^{EC} = e^{-rN}\sum_{i=0}^{N}\binom{N}{i} p^{N-i}(1-p)^i (S_N(i)-k)_+.$$

$\Pi_0^D = e^{-rN} \sum_{i=0}^{N} \binom{N}{i} p^{N-i}(1-p)^i \Pi^D_{N(i)}$

value of option at terminal time.

2.3. Derivative pricing

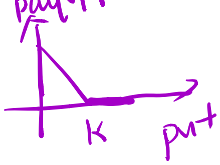
Terminal Time T.

2.3.1. Pricing European options. A European option is a derivative with exercise at expiry time. Let us assume we want to price a derivative H, whose value at scenario i is equal to $H(i)$. For instance, a European call and put with strike K are defined by $\max\{0, S_T - K\}$ call and $\max\{0, K - S_T\}$, put respectively. The price of an option can be found in a backward induction: at each specific node we find the discounted expected price of the nodes connected to that one.

convex function in terms of S

payoff

call payoff.



First for scenario i let $f_{iN} = H(i)$; for instance, for a call and a put option at time T, respectively, we have $f_{iN} = \max\{0, S_T(i) - K\}$ and $f_{iN} = \max\{0, K - S_T(i)\}$.

(1) Then using the binomial tree we have the value of the contingent claim at period j and scenario i is given by,

$$f_{ij} = e^{-r\delta t}[p f_{i,j+1} + (1-p) f_{i+1,j+1}]$$

(2) Then using the trinomial tree we have the value of the contingent claim at period j and scenario i is given by,

$$f_{ij} = e^{-r\delta t}[p_u f_{i,j+1} + p_m f_{i+1,j+1} + p_d f_{i+2,j+1}]$$

2.3.2. Pricing American options. An American option is a derivative with early exercise. This early exercise occurs if the future expected value is not higher than the current price (expected value with risk neutral). Let us assume we want to price a derivative H with early exercise, whose value at period j at scenario i is equal to $H(i,j)$. The price of an option can be found in a backward induction: at each specific node we find the discounted expected price of the nodes connected to that one and compare it with the value of the option at the same node. We take the maximum as the price of the option in that node.

First let,

$$f_{iN} = H(i, N).$$

For instance, for a call and a put option at time T, respectively, we have $f_{iN} = \max\{0, S_T(i) - K\}$ and $f_{iN} = \max\{0, K - S_T(i)\}$.

For binomial tree we have

$$\pi_j^P(i) = f_{i,j} = \max\left\{H(i,j), e^{-r\delta t}\left(p\overset{\pi^P_{j+1}(i)}{\underline{f_{i,j+1}}} + (1-p)\overset{\pi^P_{j+1}(i+1)}{\underline{f_{i+1,j+1}}}\right)\right\} \Rightarrow \text{price of the European option.}$$

For instance, for call and put option we respectively have

$$f_{i,j} = \max\left\{(S_j(i) - K)_+, e^{-r\delta t}\left(pf_{i,j+1} + (1-p)f_{i+1,j+1}\right)\right\},$$

$$f_{i,j} = \max\left\{(K - S_j(i))_+, e^{-r\delta t}\left(pf_{i,j+1} + (1-p)f_{i+1,j+1}\right)\right\}.$$

For trinomial tree we have = $\boxed{\text{European type}}$ = (下面的式的没有 "max", $D_j(i)$)

American Type = $$\pi_j^P(i) = f_{ij} = e^{-r\delta t}\max\left\{\overset{D_j(i)}{\underline{H(i,j)}}, e^{-r\delta t}\left(p_u\underset{\pi^P_{j+1}(i)}{\underline{f_{i,j+1}}} + p_m\underset{\pi^P_{j+1}(i+1)}{\underline{f_{i+1,j+1}}} + p_d\underset{\pi^P_{j+1}(i+2)}{f_{i+2,j+1}}\right)\right\}.$$

For instance, for call and put option we respectively have

$$f_{i,j} = \max\left\{(S_j(i) - K)_+, e^{-r\delta t}\left(p_uf_{i,j+1} + p_mf_{i+1,j+1} + p_df_{i+2,j+1}\right)\right\},$$

$$f_{i,j} = \max\left\{(K - S_j(i))_+, e^{-r\delta t}\left(p_uf_{i,j+1} + p_mf_{i+1,j+1} + p_df_{i+2,j+1}\right)\right\}.$$

Note that always American call price is equal to European call price, and early exercise for American call wont happen. However, the price of American put is always larger than or equal to the price of European put.

2.4. Continuous time approximations with moment matching

In the continuous time model in addition to parameter r, we have another parameter (volatility) which shows how volatile the price of asset is. In this lecture we use a Black-Scholes model that can be defined by giving the distribution of the logarithm of the division of the asset value at the end points of an interval (t_1, t_2): $X_t = \log(S_t)$, $dX_t = X_{t+\delta t} - X_t$.

$$\forall t_1 < t_2, \ln\left(\frac{S_{t_2}}{S_{t_1}}\right) \sim N\left(\left(r - \frac{\sigma^2}{2}\right)(t_2 - t_1), \sigma^2(t_2 - t_1)\right).$$

Reminder. If X is a random variable so that $\ln(X) \sim N(\mu, \sigma)$ then,

$$E(X) = e^{\mu + \frac{1}{2}\sigma^2},$$

$$Var(X) = e^{2\left(\mu + \frac{1}{2}\sigma^2\right)}\left(e^{\sigma^2} - 1\right).$$

σ = how volatility the price is

δt, 将 $[0,T]$ 分割为 small pieces

of length δt.

Now we want to approximate the continuous time model with a discrete time model. It happen to be more convenient to work with the following process and its increments

(关于这个的证明在P75)

$$X_t = \ln S_t.$$

$X_t = \log(S_t)$

Let's we split an interval $[0, T]$ to small pieces of length δt, and use the binomial model to approximate it. In this lecture notes, a good approximation happens when we can match the moments for the two models, for instance, the mean and the variance. From the definition we know that a small increment of X_t, denoted by dX_t can be found as follows

$$dX_t \approx X_{t+\delta t} - X_t = \ln S_{t+\delta t} - \ln S_t = \log\left(\frac{S_{t+\delta t}}{S_t}\right).$$

↳ small increase in X_t

But this implies that

$$dX_t \sim N\left(\left(r - \frac{\sigma^2}{2}\right)\delta t, \sigma^2 \delta t\right).$$

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

$S_t = S_0 e^{(r-\frac{\sigma^2}{2})t + \sigma W_t}$ ← Geometric Brownian motion.

So we have

$$E\left(dX_t\right) = (r - \frac{\sigma^2}{2})\delta t,$$

$= d\, S_0 e^{(r-\frac{\sigma^2}{2}) + \sigma\sqrt{t}\, Z}$, $Z \sim N(0,1)$

and

$$E\left(\left(dX_t\right)^2\right) = Var\left(dX_t\right) = \sigma^2 \delta t + \left(r - \frac{\sigma^2}{2}\right)\delta t$$

option pricing in (B-S) model

If we denote $X_t = \ln S_t$ then,

$$dX_t = X_{t+\delta t} - X_t = \ln S_{t+\delta t} - \ln S_t = \log\left(\frac{S_{t+\delta t}}{S_t}\right)$$

So therefore, we have

$$dX_t \sim N\left(\left(r - \frac{\sigma^2}{2}\right)\delta t, \sigma^2 \delta t\right).$$

2.4.1. Binomial moment matching. Now after knowing the continuous time model for small increments, let's consider a binomial model explained earlier with $u = e^{\epsilon}$ and $d = 1/u = e^{-\epsilon}$. That means our model is

$ud = 1$

$$S_{t+\delta t} = \begin{cases} uS_t & p \\ dS_t & 1-p \end{cases} = \begin{cases} S_t e^{\epsilon} & p \\ S_t e^{-\epsilon} & 1-p \end{cases}.$$

$$S_t \begin{cases} \xrightarrow{p} S_{t+\delta t} = S_t u \\ \xrightarrow{1-p} S_{t+\delta t} = S_t d. \end{cases}$$

$dX_t = \log\left(\frac{S_{t+\delta t}}{S_t}\right) = \Big\}\ -\epsilon \quad 1-p$

Let again consider $X_t = \log(S_t)$. Then one can see that $dX_t = \ln S_{t+\delta t} - \ln S_t = \begin{cases} \epsilon & p \\ -\epsilon & 1-p \end{cases}$. So we can easily find the mean and variance of dX_t form the binomial model: First Time Moment Matching:

$$E(dX_t) = p\epsilon + (1-p)(-\epsilon) = (2p-1)\,\epsilon; = \left(r - \tfrac{1}{2}\sigma^2\right)\delta t \leftarrow \text{First moment}$$

$$E\left((dX_t)^2\right) = p\epsilon^2 + (1-p)\,\epsilon^2 = \epsilon^2. = \sigma^2\delta t + \left(\left(r - \tfrac{\sigma^2}{2}\right)\delta t\right)^2 \leftarrow \text{Second moment}$$

On the other hand, we know these quantities from continuous process. Now, in order to approximate the continuous time model with a binomial tree we match the first and the second moments of dX_t as follows:

$$\begin{cases} \underbrace{(2p-1)\,\epsilon}_{\text{binomial model}} = E(dX_t) = \underbrace{\left(r - \frac{\sigma^2}{2}\right)\delta t}_{\text{continuous model}} \\ \underbrace{\epsilon^2}_{\text{binomial model}} = E\left((dX_t)^2\right) = \underbrace{\sigma^2\delta t + \left(\left(r - \frac{\sigma^2}{2}\right)\delta t\right)^2}_{\text{continuous model}}. \end{cases} \Big\}\ 求\ \epsilon\ 和\ p.$$

From the second equation,

$$\epsilon = \sqrt{\sigma^2\delta t + \left(\left(r - \frac{\sigma^2}{2}\right)\delta t\right)^2},$$

(put in 1st moment equation to obtain...).

If we use it in the first equation we get

$$p = \frac{\left(r - \frac{\sigma^2}{2}\right)\delta t}{2\epsilon} + \frac{1}{2}.$$

(*) we do option pricing in the Binomial Model with parameters $(r, u = e^{\epsilon}, d = e^{-\epsilon}, p)$ where p is given by (*) to rounded a approximation of option price in the continous model.

EXAMPLE 37. Let us assume a true continuous model S_t for a stock price is $\ln\left(\frac{S_{t+\delta t}}{S_t}\right) \sim N\left(\left(r - \frac{\sigma^2}{2}\right)\delta t, \sigma^2\delta t\right)$. Using a binomial model by matching the first and the second moment, where $\delta t = 1$, $dX_t = \begin{cases} 0.3 & p_u \\ -0.3 & p_d \end{cases}$, $r = 0.045$, $\sigma = 0.3$, we want to find the price of an American option for expiration $T = 2$. We also have step (1) $\epsilon = \sqrt{\sigma^2\delta t + \left(\left(r - \frac{\sigma^2}{2}\right)\delta t\right)^2} = \sqrt{0.09 \times 1 + \left(\left(0.045 - \frac{0.09}{2}\right) \times 1\right)^2} = 0.3$. So we have to find the first

Using moment matching give an approximation of the price of European call option with values $S_0 = K$ by usi Binomial Model.

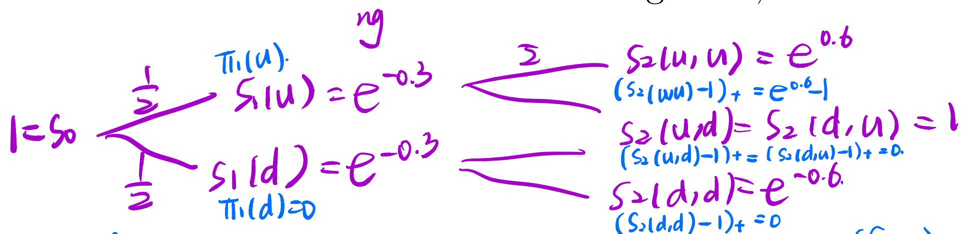
and the second moments:

$$\underbrace{0.3p_u - 0.3p_d}_{} = E(dX) = E\left(\ln\left(\frac{S_{t+\delta t}}{S_t}\right)\right)$$
$$= \left(r - \frac{\sigma^2}{2}\right)\delta t = 0.045 - \frac{0.09}{2} = 0,$$

or $p_u - p_d = 0$.Also we know $p_u + p_d = 1$. So we have

$$\begin{cases} p_u - p_d = 0 \\ p_u + p_d = 1 \end{cases}.$$

Summing up first and second equations we get $p_d = p_u = 0.5$. Therefore, we have to consider the following trees,



The Payoff of European call option with strike $k=1$ is given by $(S_2-1)_+$ // $\Pi_1(d) = e^{-r\delta t}\left(\frac{1}{2}\Pi_2(u,d) + \frac{1}{2}\Pi_2(d,d)\right)$.

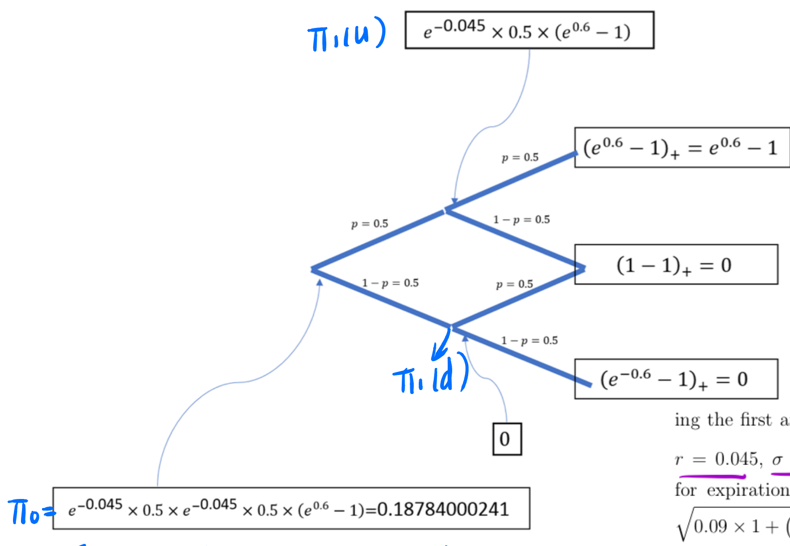
$\Pi_1(u) = \left(e^{-0.045}\right)\left(\left(\frac{1}{2}\right)(e^{0.6}-1) + \frac{1}{2}(0)\right)$

f_{10} $= \left(e^{-0.045}\right)(0.5)(e^{0.6}-1)$

$\Pi_0 = e^{-r\delta t}\left(\frac{1}{2}\Pi_1(u) + \frac{1}{2}\Pi_1(d)\right)$

$= \left(e^{-0.045}\right)^2 (0.5)^2 (e^{0.6}-1)$

Tree for option price:



≈ European call option with strike K=1 in (B-S) model

ing the first and the second moment, where $\delta t = 1$, $dX_t = \begin{cases} \ldots & p_u \\ -0.3 & p_d \end{cases}$, option pr. in the com model. $r = 0.045$, $\sigma = 0.3$, we want to find the price of an American option for expiration $T = 2$. We also have step①: $\epsilon = \sqrt{\sigma^2 \delta t + \left(\left(r - \frac{\sigma^2}{2}\right)\delta t\right)^2} = \sqrt{0.09 \times 1 + \left(\left(0.045 - \frac{0.09}{2}\right) \times 1\right)^2} = 0.3$. So we have to find the first

Using moment matching give an approximation of the price of European call option with values $S_0 = 1 = K$ by use Binomial Model.

FIGURE 2.4.1. European call.

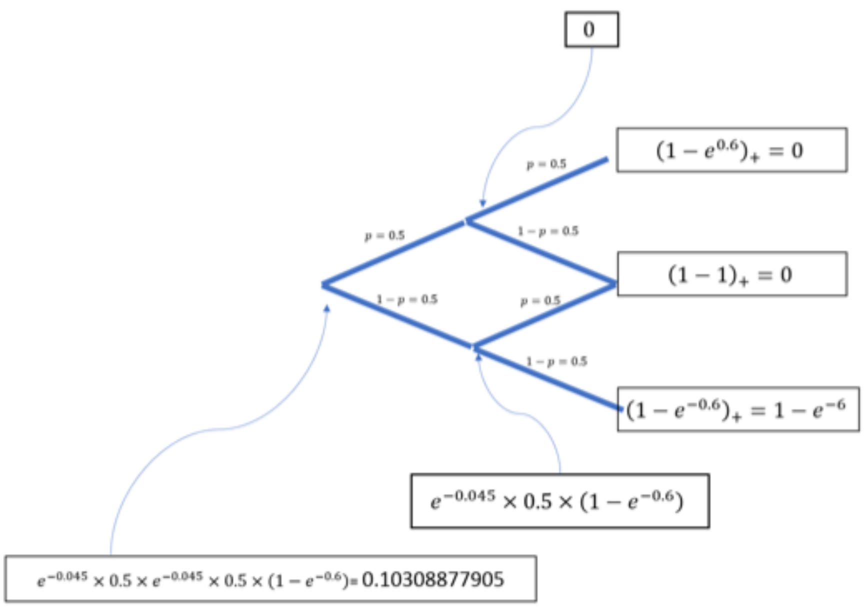


FIGURE 2.4.2. European put.

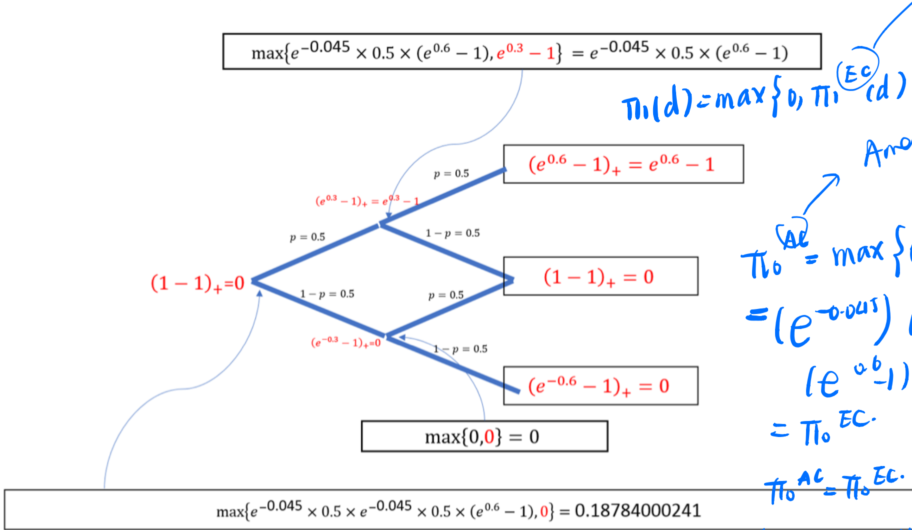
$S_2(u,u) = e^{0.6}$

$S_2(u,u) = e^{0.6}$

$1 = S_0$, with probability $\frac{1}{2}$: $S_1(u) = e^{0.3}$; with probability $\frac{1}{2}$: $S_1(d) = e^{-0.3}$

$S_2(u,d) = S_2(d,u) = 1$

$S_2(d,d) = S_2(d,d) = e^{-0.6}$



European call

$\Pi_1(d) = \max\{0, \Pi_1^{(EC)}(d)\} = \max\{0,0\}.$

American

$\Pi_0^{AC} = \max\{0, \frac{1}{2}\Pi_0^{EC} e^{-0.045}\}.$

$= (e^{-0.045})(0.5)(e^{-0.045})(0.5)(e^{0.6}-1)$

$= \Pi_0^{EC}.$

$\Pi_0^{AC} = \Pi_0^{EC} \approx$ Price of the American call option with striking $k=1$ in the continuous time (B-S) model.

FIGURE 2.4.3. American call.

Tree for option prices:

$\Pi_0^{AC} = \Pi_0^{EC}$, with $\frac{1}{2}$: $\Pi_1^{AC}(u) = \Pi_1^{EC}(u)$; with $\frac{1}{2}$: $\Pi_1^{AC}(d) = 0$

$\Pi_2^{AC}(u,u) = \Pi_2^{EC}(u,u) = e^{0.6}-1$

$\Pi_2^{AC}(u,d) = \Pi_2^{AC}(d,u) = 0 = \Pi_2^{EC}$

$\Pi_2^{AC}(d,d) = \Pi_2^{EC}(d,d) = 0$

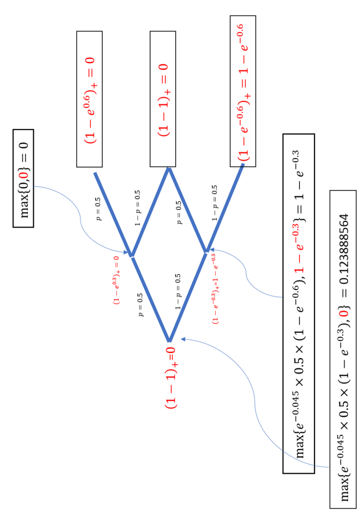


FIGURE 2.4.4. American put.

2.4.2. Trinomial moment matching. Like before, let dX_t be a small increment which can be approximated by three values, $\epsilon, 0, -\epsilon$ with probabilities p_u, p_m, p_d. In this case note that the value of the asset is given by

$$dX_t = \log\left(\frac{S_{t+\delta t}}{S_t}\right) = \ln S_{t+\delta t} - \ln S_t = dX_t = \begin{cases} \epsilon & p_u \quad \log(S_t e^{\epsilon}/S_t) \\ 0 & p_m \, . \quad \log\left(\frac{S_t}{S_t}\right) \\ -\epsilon & p_d \quad \log\left(\frac{S_t e^{-\epsilon}}{S_t}\right) \end{cases}$$

This is associated with the model $S_{t+\delta t} = \begin{cases} S_t e^{\epsilon} & p_u \\ S_t & p_m \\ S_t e^{-\epsilon} & p_d \end{cases}$, where $u = e^{\epsilon}$and $d = 1/u = e^{-\epsilon}$. Let's see the moment from the discrete model:

$$E(dX) = p_u\epsilon + p_m \times 0 - p_d\epsilon = p_u\epsilon - p_d\epsilon = \left(r - \frac{\sigma^2}{2}\right)\delta t$$

$$E\left((dX)^2\right) = p_u\epsilon^2 + p_m \times 0^2 + p_d\epsilon^2 = p_u\epsilon^2 + p_d\epsilon^2 = \sigma^2\delta t + \left(\left(r - \frac{\sigma^2}{2}\right)\delta t\right)^2.$$

Now, in order to approximate the continuous time model with a trinomial model we match the first and second moments of dX_t. Matching the expectation and variance, we get

There are 3 unknown parameters, with only 2 equation. → (in general) infinitly many solutions.

$$\begin{cases} p_u\epsilon - p_d\epsilon = \delta t(r - \frac{\sigma^2}{2}) = v\delta t \\ p_u\epsilon^2 + p_d\epsilon^2 = \sigma^2\delta t + v^2(\delta t)^2 \\ p_u + p_m + p_d = 1 \end{cases},$$

① Fix one of the parameters ⇒ $\epsilon = 3\sqrt{\delta t}$
② matching one more higher moment

where $v = r - \frac{\sigma^2}{2}$. The last condition is simply because p_d,p_m and p_u constitute a probability measure. After solving the equations we get

$P_u + P_m + P_d = 1$

plug in equation $\epsilon = 3\sqrt{\delta t}$, we

$$p_u = \frac{1}{2}\left(\frac{\sigma^2\delta t + v^2(\delta t)^2}{\epsilon^2} + \frac{v\delta t}{\epsilon}\right),$$

$$p_m = 1 - \left(\frac{\sigma^2\delta t + v^2(\delta t)^2}{\epsilon^2}\right),$$

$$p_d = \frac{1}{2}\left(\frac{\sigma^2\delta t + v^2(\delta t)^2}{\epsilon^2} - \frac{v\delta t}{\epsilon}\right).$$

The extra parameter ϵ, which gives us one degree of freedom, is set to be $\epsilon = 3\sqrt{\delta t}$.

Example 38. Let us assume a true continuous model S_t for a stock price is $\ln\left(\frac{S_{t+\delta t}}{S_t}\right) \sim N\left(\left(r-\frac{\sigma^2}{2}\right)\delta t, \sigma^2\delta t\right)$.Using a trinomial model by matching the first and the second moment, where $\delta t = 1$, $dX = \begin{cases} 3\sqrt{\delta t} = 3, & p_u \\ 0, & p_m \\ -3\sqrt{\delta t} = -3, & p_d \end{cases}$, $r = 0.045$, $\sigma = 0.3$, we want to find the price of an American option for expiration $T = 2$. So we have to find the first and the second moments:

① First moment:

$$\begin{aligned} 3p_u - 3p_d = E(dX) &= E\left(\ln\left(\frac{S_{t+\delta t}}{S_t}\right)\right) \\ &= \underbrace{\left(r-\frac{\sigma^2}{2}\right)}_{=0}\delta t = 0.045 - \frac{0.09}{2} = 0, \end{aligned}$$

or $p_u - p_d = 0$. And we have

② second moment:

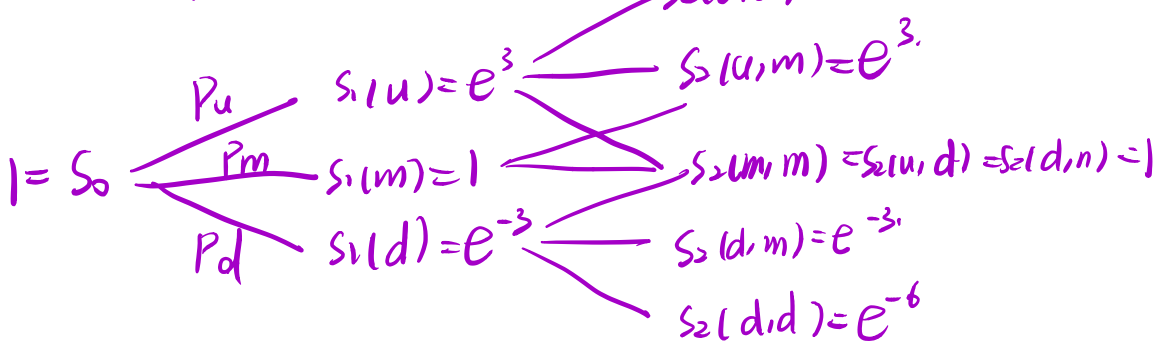
$$\begin{aligned} 9p_u + 9p_d &= E\left((dX)^2\right) \\ &= E\left(\left(\ln\left(\frac{S_{t+\delta t}}{S_t}\right)\right)^2\right) = \sigma^2\delta t + \left(\left(r-\frac{\sigma^2}{2}\right)\delta t\right)^2 = 0.09. \end{aligned}$$

Also we know $p_u + p_m + p_d = 1$. So we have

$$\begin{cases} p_u - p_d = 0 \\ p_u + p_d = 0.01 \\ p_u + p_m + p_d = 1 \end{cases}.$$

Summing up first and second equations we get $2p_u = 0.01$, or $p_u = 0.005$. So $p_d = p_u = 0.005$ and $p_d = 0.99$. Therefore, we have to consider the following trees,

Tree of Price.



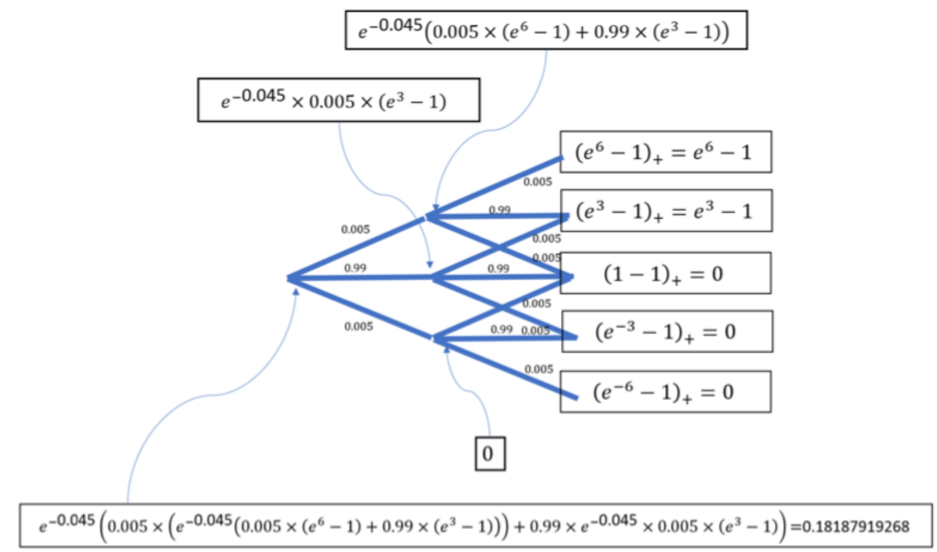


FIGURE 2.4.5. European call.

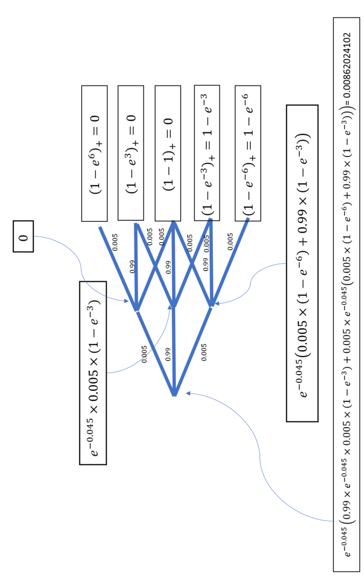


FIGURE 2.4.6. European put.

$\Pi_1^{AC}(d) = \max\{(S_1(d)-1)_+, e^{-r\delta t}(P_d\Pi_2^{AC}(d,d) + P_m\Pi_2^{AC}(d,m) + P_n\Pi_2(m,n))\} = 0.$

$\Pi_1^{AC}(m) = \max\{(S_1(m)-1)_+, e^{-r\delta t}(P_n\underbrace{\Pi_2^{AC}(m,u)}_{e^3-1})\}.$

$= (e^{-0.045})(0.005)(e^3-1)$

$\Pi_1^{AC}(u) = \max\{(S_1(u)-1)_+, e^{-r\delta t}(P_m\underbrace{\Pi_2^{AC}(u,m)}_{(e^3-1)} + P_u\underbrace{\Pi_2^{AC}(u,u)}_{(e^6-1)})$

$= \max\{(e^3-1), e^{-0.045}((0.99)(e^3-1) + (0.05)(e^6-1)\}.$

$\Pi_0^{AC} = \max\{\cancel{(S_0-1)_+}, e^{-r\delta t}(P_n\Pi_1(u) + P_m\Pi_1^{AC}(m) + P_d\Pi_1^{AC}(d))$ 0.

$= \Pi_0^{EC}.$

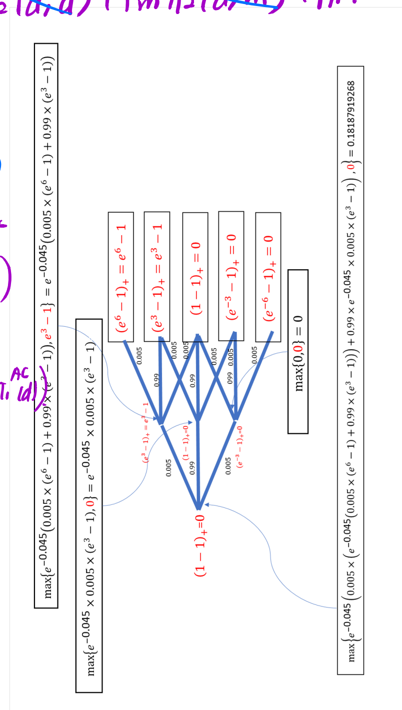


FIGURE 2.4.7. American call.

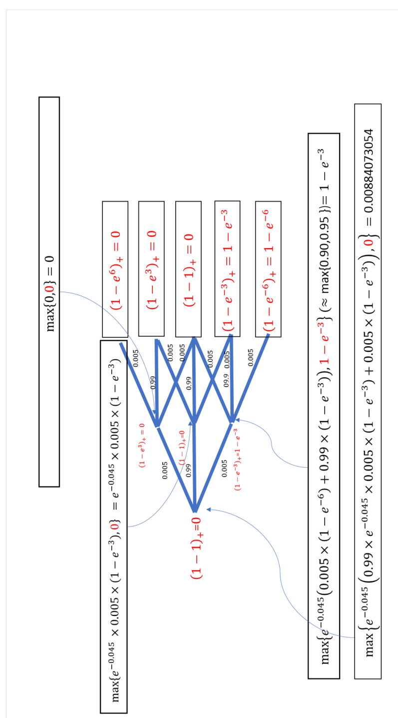


FIGURE 2.4.8. American put.

2.5. Exercise

EXERCISE 39. Assume we want to price derivatives on a stock whose value is given by a geometric Brownian motion. The value of the volatility and return are

$$\sigma = 0.1$$
$$r = 0.01$$

Now we want to use a recombining binomial tree with a time-step $\delta t = 1$ and horizon time $2\delta t$

Assume the price of the stock at time 0 is $S_0 = 10$ and stick price $K = 10$.

a) First use a binomial tree for approximating the price of a European call option. Draw the trees for the stock and for the option, and clearly assign the values to the nodes (Use a big graph please).

b) Second use the binomial three to find the price of an America put option with the same strike price. Draw the tree for the option, and clearly assign the values to the nodes (Use a big graph please).

c) Compare the two answers and tell which one is greater.

EXERCISE 40. Assume we want to price derivatives on a stock whose value is modeled by a geometric Brownian motion with volatility and risk free rate given as

$$\sigma = 0.2,$$
$$r = 0.02.$$

We assume the physical probability measure is equal to the risk neutral probability measure. Now we want to estimate the value of a European and American call options by using binomial and trinomial trees with a time-step $\delta t = 1$ and horizon time $2\delta t$

Assume the price of the stock at time 0 is $S_0 = 1$ and the strike value is $K = 1$ which expires at $T = 2$.

a) Use a recombining binomial tree as an approximation. Draw the trees for the stock and for the option, and clearly assign the values to the nodes (Use a big graph please).

b) Now assume we want to price an American call option with the same K and T. Draw the trees for the option, and clearly assign the values to the nodes (Use a big graph please). Compare your answer to your finding in part a). Which one is greater and why?

c) Use a recombining trinomial tree as an approximation. Draw the trees for the stock and for the option, and clearly assign the values to the nodes (Use a big graph please). Note that,

$$\begin{aligned}
&u = \exp(\epsilon), && d = 1/u \\
&p_u = (1/2)\left(\frac{\sigma^2\delta t+\nu^2(\delta t)^2}{\epsilon^2} + \frac{\nu\delta t}{\epsilon}\right), && p_d = (1/2)\left(\frac{\sigma^2\delta t+\nu^2(\delta t)^2}{\epsilon^2} - \frac{\nu\delta t}{\epsilon}\right) \\
&p_m = 1 - p_u - p_d, && \epsilon = 3\sqrt{\delta t} \\
&\nu = r - \frac{\sigma^2}{2}, && \text{discount } = \exp(-r\delta t)
\end{aligned}$$

d) Compare the answers from part a) and c) and tell which one, in general, is a better answer.

EXERCISE 41. Consider a re-combining binomial tree for two steps $N = 2$. Assume that the strike price is equal to the stock price at time zero i.e., $K = S_0$. Assume that the interest rate is equal to zero i.e., $r = 0$ and that at each node the stock price moves up by a specific factor $u = 2$ or down by $d = 1/2$.

a)If a European put option price is equal to $1/12$, what is the value of S_0 (you have to draw the binomial tree for the put option).

b) Given the value that you have found in part a), find the price of an American put option with the same stick price (you have to draw the binomial tree of the stock and the American option).

c) Compare the American put option and the European put option prices. Which one is larger? Did you expect this?

d) The following program is a program to price a European put price.

Please read and understand it. What is the necessary change for a program which prices an American option (Note q in the following is the probability of going up and 1-q the probability of going down. In addition S_0=s).

```
for i=1:N+1
    f(N+1,i)=max{K-u(N-i+1)*d(i-1)*s,0};
end
for i=1:N
    for j=1:N-i+1
        f(N-i+1,j)=q*f(N-i+2,j)+(1-q)*f(N-i+2,j+1);
    end
end
```

EXERCISE 42. Assume we want to price derivatives on a stock whose value is given by a geometric Brownian motion. The value of the volatility and return are

$$\sigma = 0.1$$
$$r = 0.01$$

Now we want to use a tree with a time-step $\delta t = 1$ and horizon time $2\delta t$. Assume the price of the stock at time 0 is $S_0 = 10$.

a) First, use a binomial tree for approximating the prices of European call and put European options on the asset with strike price $K = 10$. Draw the trees for the stock and for the options, and clearly assign the values to the nodes.

b) Second, use the binomial three to find the price of an America put option with the same strike price. Draw the tree for the option, and clearly assign the values to the nodes.

c) Compare the two answers to the prices of European put option and American put option in step a) and b) and tell which one is greater. Justify your answer.

EXERCISE 43. Let us assume a true continuous model S_t for a stock price is $\ln\left(\frac{S_{t+\delta t}}{S_t}\right) \sim N\left(\left(r - \frac{\sigma^2}{2}\right)\delta t, \sigma^2\delta t\right)$. Let's assume $r = \frac{\sigma^2}{2}$, $\delta t = 1$ and $T > 0$.

a) Using a binomial model by matching the first and the second moment, find the price of an American option for expiration T.

b) Using a trinomial model by matching the first and the second moment, find the price of an American option for expiration T.

Consider Stochastic SDE for the stock price in (B-S) market

$dS_t = S_t r\,dt + S_t \sigma\, dW_t$

Solution of this SDE:

$S_t = S_0 e^{(r-\frac{1}{2}\sigma^2)t + \sigma W_t}$, where $W_t \sim N(0,t)$

$\stackrel{d}{=} S_0 e^{(r-\frac{1}{2}\sigma^2)t + \sigma\sqrt{t}Z}$, where $Z \sim N(0,t)$

So, one dimensional marginal distribution of stock price are log normal

Simulation in Finance

Motivation

Generally Speaking, in mathematical finance, the price of European-type options are given by mathematical expectation of some "complex" random variables:

For example, in the celebrated B-S market model

$B_t = B_0 e^{rt}$; $t \in [0,T]$

$S_t = S_0 e^{(r-\frac{1}{2}\sigma^2)t + \sigma W_t}$, where $W_t \sim N(0,t)$

Geometric Brownian motion.

CHAPTER 3

Monte-Carlo simulation in finance

3.1. Introduction

In finance, in many cases, we need to find average for a particular random variable. For example, in a continuous time model that was introduced in the previous chapter, the price of an option can be found by a simple expectation.

(a) European call option with strike price k; and payoff $(S_T-k)_+$.

Call:

$= e^{-rT} E_Q[(S_T-K)_+]$, where $W_T \sim N(0,T)$.

$$C(K,T,S_0) = S_0 e^{-rT} E\left[\left(e^{(r-\frac{1}{2}\sigma^2)T+\sigma W_T} - K\right)_+\right].$$

Put:

$= e^{-rT} E_Q[(K-S_T)_+]$.

$$P(K,T,S_0) = S_0 e^{-rT} E\left[\left(K - e^{(r-\frac{1}{2}\sigma^2)T+\sigma W_T}\right)_+\right].$$

Main Question: How to compute the Expectation?

Mathematical Expectation. ⟷ Emprical Mean Sample Mean.

where $W_T \sim N(0,T)$. If we introduce a sample,

$$S_{i,T} = S_0 e^{(r-\frac{1}{2}\sigma^2)T+\sigma W_{i,T}}, i = 1, ..., n,$$

Let's take a sample of $N(0,T)$, namely that $(N_i, i \geq 1)$ are the outcomes of independent Gaussian distribution with mean 0 and variance T; and. Instead we look to the sample mean; i.e.: (Sample sign N)

where $W_{i,T}, i = 1, ..., n$ is an i.d.d. sample of W_T with $N(0,T)$ distribution then in order to approximate the price of the European call and put, one can use the empirical means,

$$C(T,S_0,K) \approx \quad e^{-rT}E(S_T-K)_+ \approx e^{-rT}\frac{(S_{1,T}-K)_+ + ... + (S_{n,T}-K)_+}{n}$$

, where $S_T = S_0 e^{(r-\frac{1}{2}\sigma^2)T+\sigma N_i}$ $i=1,2,\dots N$.

similar for European put options:

$$e^{-rT}E(K-S_T)_+ \approx e^{-rT}\frac{(K-S_{1,T})_+ + ... + (K-S_{n,T})_+}{n}$$

But still there is no guarantee that the right hand side can approximate the left hand side (even though we feel it is correct). So we have to find expectation of a random variable.

We have two issues,

(1) How to generate a sample.

(2) How to find more accurate mean. ⇒ By Law of Large Number (LLN)

3.1.1. Random generator. A uniform random generator is a machine which gives or generates a sequence $u_1, u_2, u_3, \ldots$ of numbers which has a uniform distribution; the chance of having any number between (a, b) is equal. In mathematical terms the probability of having any number between $c, d \in (a, b)$ is,

$$\frac{d-c}{b-a}.$$

So it means a sample $u_1, u_2, u_3, \ldots$ is uniform if

$$\frac{\#\left\{u_i | c \leq u_i \leq d\right\}}{n} \approx \frac{d-c}{b-a},$$

when n is big.

Generating uniform numbers is not an easy job. There are different random generator

(1) Based on physical experiments.

(2) Based on mathematical and computational methods .

(3) Based on human experiments.

The most used one is one generates by mathematical methods. For instance the most known method is called the linear congruential generator that generates number by,

$$u_{n+1} \equiv^m a u_n + c,$$

for different choices of a and c. These numbers can be found on trial and error and can be different for different generators. For instance, in Borland C/C++, the numbers are $m = 2^{32}$, $a = 22695477$ and $c = 1$.

Definition 44. A (general) **sample** is a sequence $(x_i)_{i=1}^n$ that are generated independently and has a particular distribution F i.e.,

$$F(x) = \lim_{n \to \infty} \frac{\#\{x_i | x_i \leq x\}}{n}.$$

3.1.2. Law of Large Number (LLN). As was mentioned earlier, we want to use an empirical average to estimate the real average. The following theorem can make us sure such approximation is legitimate under some mild conditions.

Def: [Different Type of convergence in probability theory]

Let $(X_n; n \geq 1)$ be a sequence of R.V. Then

(a) Almost sure convergence (a.s.)

We say $X_n \xrightarrow[\text{as } n\to\infty]{a.s.} X$ if $P(\omega \in \Omega ; \lim_{n\to\infty} X_n(\omega) = X(\omega)) = 1$

↳ random objective.

sequence of number converge to this number.

(b) convergence in probability (P)

we say $X_n \xrightarrow[\text{as } n\to\infty]{P} X$ if, $\forall \varepsilon > 0$.

$$\lim_{n\to\infty} P(\omega ; |X_n(\omega) - X(\omega)| \geq \varepsilon) = 0$$

limit outside the property equal to 0.

(c) L^P- Convergence ; $P \geq 1$

↳ a number.

We say $X_n \xrightarrow[\text{as } n\to\infty]{L^P} X$ if $E|X_n - X|^P \to 0$.

(d) "Weak Convergence" (convergence in distribution)

We say $X_n \xrightarrow[\text{as } n\to\infty]{D} X$ if

$$F_{X_n}(x_0) = P(X_n \leq x_0) \xrightarrow{\text{converge to}} F_X(x_0) = P(X \leq x_0)$$

$\forall x_0 \in \mathbb{R}$ such that $x \longmapsto F_X(x)$ is continuous at x_0

distribution function of the limit.

Remark:

$$\begin{array}{c} a.s. \\ \Downarrow \text{ implies.} \\ P \Rightarrow D. \\ \Uparrow \text{ implies.} \\ L^P \end{array}$$

(which one is weaker / stronger)

Go back to the LLN.

Thm: Let X be a real-valued, r.v., and $X_1, X_2, \cdots$ be an independent copies of X; meaning that X_i are independent and $X_i \sim X$, $\forall i \geq 1$

have the same distribution.

Let $\overline{X}_n := \dfrac{X_1 + X_2 + \cdots X_n}{n}$. a new random objective.

(a) (Weak form) suppose that $E|X| < \infty$. Then

$$\overline{X}_n \xrightarrow[\text{converge}]{P} \mu (=: E(X)) \quad \text{[average]}$$

first deviation. [don't need second moment to be finite].

(b) (strong form) suppose that $E(X) < \infty$. Then

$$\overline{X}_n \xrightarrow{a.s.} \mu (=: E(X))$$

Remark: For (a) and (b) to hold, we do not need $E|X|^2 < \infty$. → second moment

first 可证明它

proofs: [under an extra (extremely demanding) assumption that $E[X]^4 < \infty$] → more precise statement

$\overline{X}_n - \mu = \frac{1}{n} \left(\sum_{i=1}^{n} (X_i - \mu) \right)$ This implies that mean of $(X_i - \mu) = 0$

⇒ 所以只有2个 contribution

try to show it converge

$$E\left((\overline{X}_n - \mu)^4 \right) = \frac{1}{n^4} E\left(\sum_{i=1}^{n} (X_i - \mu) \right)^4.$$

↳ indep. to each other 因为 independence $E((X-\mu)^2))^2$. so.

→ all X_i has same distribution.

same distribution but indp.

by Binomial formula. ⇒

$$= \frac{1}{n^4} \left(nE(X-\mu)^4 + 3n(n-1) E\left((X-\mu)^2 (X'-\mu)^2 \right) \right)$$

↳ one factor and another factor

$X' \overset{\perp}{\sim} X$ has same distribution and independent.

$$\leq \frac{1}{n^4} \{ n^2 k \} \text{ when } k = 3 \max\{ E(X-\mu)^4, (E(X-\mu)^2)^2 \}$$

forth-moment second moment.

$$= \frac{k}{n^2} \to 0 \text{ as } n \to \infty$$

Moreovers: $E\left(\sum_{n\geq 1} (\overline{X}_n - \mu)^4 \right) = \sum_{n \geq 1} E(\overline{X}_n - \mu)^4 \leq \sum_{n\geq 1} \frac{k}{n^2} < \infty$.

$\Rightarrow \sum_{n \geq 1} (\overline{X}_n - \mu)^4 < \infty$, a.s $\Rightarrow (\overline{X}_n - \mu)^4 \to 0$. as $n \to \infty$ a.s. $\Rightarrow \overline{X}_n \to \mu$, a.s.

Summary

Q: $F(x)$ Mathmatical Expecatation → empiral average sampling mean

Q(1) what is a sample of a given distribution? It is easy to generate a sample?

Q(2) If $(X_n; n\geq 1)$ be a sample of distribution of X, then we expect that

$$E(X) \approx \frac{X_1 + \cdots X_n}{n} \text{ for large } n \in \mathbb{N}$$

Back to Q2: Law of Large Number (LLN)

(short form) Let X be any distribution, with $E(X) < \infty$ (so, $\mu := E(X) < \infty$). Let $(X_n; n\geq 1)$ be an independent copies of X.

Then $\bar{X}_n := \frac{X_1 + \cdots X_n}{n} \xrightarrow[\text{as } n\to\infty]{a.c.} \mu$

I.e. $\forall w \in \Omega$

$$\bar{X}_n(w) = \frac{X_1(w) + \cdots X_n(w)}{n} \to \mu \text{ as } n\to\infty.$$

Proof: based on extra assumption than $E|X^n| < \infty$

Let's back to (Q1):

Def = [Sample]

Let's F be a probability distribution. We say a sequence $(X_n; n\geq 1)$ is a sample at distribution F if the following hold:

(a) They are generated independently.

(b) More importantly; $\forall x \in \mathbb{R}$

$$\lim_{n\to\infty} \frac{\#\{X_i ; X_i \leq x, i \leq n\}}{n} = F(x)$$

You may also think of a sample of a given distribution as a realization of independent copies of the given distribution.

How to generate a sample of a given probability distribution?

(A) The Inversion Method:

Def. [Uniform Distribution]

Let an interval $[a,b] \subseteq \mathbb{R}$. we say a random variable X follow a uniform distribution on interval $[a,b]$, and we write $X \sim U(a,b)$, if

$$F_X(x) = \begin{cases} 0 & \text{if } x \leq a \\ \frac{x-a}{b-a} & \text{if } a \leq x \leq b \\ 1 & \text{if } x \geq b \end{cases} \qquad f_X(x) = \begin{cases} \frac{1}{b-a} & \text{if } x \in (a,b) \\ 0 & \text{otherwise} \end{cases}$$

Thm: Let X be a R.V with probability distribution function F (F_X), and that F is continuous. Let $U \sim U(0,1)$, then. $Y := F^{-1}(U) \sim X$ ($Y =_d X$)

[In other words. Uniform distribution serves as a mother distribution to generate any probability distribution].

Proof: Increasing, and continuity assumption impose on F yield that F^{-1} exist as an ordinary inverse function. So random variable Y is well-defined.

Next, we want to show that, $\forall x \in \mathbb{R}$

$$P(Y \le x) = P(X \le x) = F(x)$$

Now, note that

$$P(Y \le x) = P(F^{-1}(U) \le x) = P(\overbrace{F(F^{-1}(U))}^{U} \le F(x))$$

$$= P(U \le F(x)) = F(x)$$

THEOREM 45. *(Law of Large Number) Let $X_1, X_2, \ldots$ be a sequence of an i.i.d random variables whose mean is μ and variance is $\sigma^2 < \infty$. (This is important, both mean and variance have to be finite). Then, Law of Large Number (LLN) guarantees that*

$$\frac{X_1 + \ldots + X_n}{n} \to \mu.$$

This guarantees that if we have a sample with finite mean and variance then the empirical mean will converge to real mean. But it is really important to check all the conditions. For instance in the following we have a case where the variance does not exists and as a result the mean is not guaranteed to be convergent. Let,

$$F(x) = \begin{cases} (x-1)/x & \text{if } \; x \geq 1 \\ 0 & o.w. \end{cases}.$$

As one can see,

$$F(x) = \frac{x-1}{x}, x \geq 1 \Longrightarrow f(x) = \frac{1}{x^2}$$

$$\Longrightarrow E(X) = \int_0^\infty x \frac{1}{x^2} dx = \infty,$$

so mean is not finite i.e., the theorem cannot be applied.

3.2. Inversion method

Inversion method is a method which allows us to generate a sample for any distribution whose cumulative distribution function, F_X, and its inverse, F_X^{-1}, are known. Therefore, we will have a quick review on some probability concepts before stating the method. Recall that for any random variable X the Cumulative Distribution Function of X (CDF) is defined as,

$$F_X(x) = P(X \leq x).$$

If this function is differentiable then the Probability Density Function (PDF) is introduced as,

$$f_X(x) = F_X'(x)$$

Therefore, we have

$$F_X(x) = \int_{-\infty}^x f_X(y) dy.$$

For example, a uniform $U(0,1)$ distribution has the following CDF and PDF:

$$F(x) = \begin{cases} 0 & \text{if} \quad x \leq 0 \\ x & \text{if} \quad 0 \leq x \leq 1 \\ 1 & \text{if} \quad x \geq 0 \end{cases}$$

and

$$f(x) = \begin{cases} 1 & \text{if} \quad 0 \leq x \leq 1 \\ 0 & \text{o.w.} \end{cases}$$

As we have seen there are several methods to generate samples for $U(0,1)$. Now, the question is how we can generate samples for other important distributions.The idea of the method is based on the following simple theorem.

Theorem 46. *If U is a random variable with $U(0,1)$ distribution then, $F_X^{-1}(U)$ is a random variable whose distribution is the same as X.*

$Y = F^{-1}(U) \sim X \; (Y =_d X)$

Proof. When F_x is invertible, $P(F_X^{-1}(U) \leq x) = P(U \leq F_X(x)) = F_U(F_X(x)) = F_X(x)$. Increasing and continuity assumptions impose on F yield that F^{-1} exists as an ordinary inverse function. So, random variable Y is well. $\square$

$P(Y \leq x) = P(X \leq x) = F(x)$, Now note that: $P(Y \leq x) = P(F^{-1}(U) \leq x)$

$= P(\underbrace{F(F^{-1}(U))}_{U} \leq F(x))$

$= P(U \leq F_X(x))$

Example 47. For example, let us consider the following CDF,

$$F_x(x) = \begin{cases} \frac{x-1}{x} & \text{if} \quad x \geq 1 \\ 0 & \text{if} \quad x \leq 1 \end{cases}$$

Then $F^{-1}(y) = \frac{1}{1-y}$. So if we generate $u_1, u_2, ..., u_n$ numbers from a $U(0,1)$ distribution, a random variable with CDF F_X will be,

$$\frac{1}{1-u_1}, \frac{1}{1-u_2}, ..., \frac{1}{1-u_n}.$$

3.3. Acceptance-Rejection method

For two main reasons we might be unable to properly generate a random sample with inversion method.

Although inversion method is simple, however, the price to pay is that we need to comput F^{-1}, which in general may not easy Job.

(1) F_X is not known (or F_X^{-1}) ;
(2) The inversion method is not efficient enough.

(b) The inversion method still works if we drop the continnuitly assumption on probability distribution function F.

In general, one can define the generalized inversion function $F^{-1}(F_g^{+})$

$F^{+}: [0,1] \longrightarrow \mathbb{R}$

$y \longrightarrow F^{+}(y) = \inf\{x : F(x) \geq y\}$

F^{+} is a well-defined, and moreover

(i) $F^{+}(F(x)) \leq x$; $\forall x \in \mathbb{R}$

(ii) $F(F^{+}(x)) \geq x$; $\forall x \in (0,1)$

cdf Inversion Method 的例题.

[Exponential Distribution]

Let $X \sim Exp(\lambda)$, when $\lambda > 0$, that is

$$f_X(x) = \lambda e^{-\lambda x}; \ x \geq 0$$

we want to apply the inversion method to generate exponential distribution out of $U(0,1)$ distribution.

(i) We need to compute $F_X(x) = P(X \leq x) = \int_0^x f_X(y)\,dy$

$$= \int_0^x \lambda e^{-\lambda y} dy = 1 - e^{-\lambda x}; \ x \geq 0$$

(ii) We need to compute F_X^{-1}

$$y = 1 - e^{-\lambda x} \Rightarrow x = -\frac{1}{\lambda}\log(1-y)$$

This gives as

$$F_X^{-1}(y) = -\frac{1}{\lambda}\log(1-y); \ \forall y \in (0,1)$$

(iii) The random variable $Y := -\frac{1}{\lambda}\log(1-U) =_d X \sim Exp(\lambda)$ and $U(0,1) \sim U$

Therefore, if $u_1, u_2, \cdots u_N$ be a sample of $U(0,1)$, then

$$-\frac{1}{\lambda}\log(1-u_1), -\frac{1}{\lambda}\log(1-u_2), \cdots, -\frac{1}{\lambda}\log(1-u_N)$$

is a sample of $Exp(\lambda)$

Since, $1-U \sim U(0,1)$, when $U \sim U(0,1)$, hence

$-\frac{1}{\lambda}\log(u_1), \cdots, -\frac{1}{\lambda}\log(u_N)$ is also a sample of size N for $Exp(\lambda)$ with less computional operation.

Ex: [Weinbull Distribution]

$X \sim Weinbull(\alpha, \beta)$, $\alpha, \beta > 0$, if

$$f_X(x) = \frac{\alpha}{\beta} x^{\alpha-1} e^{-\frac{x^\alpha}{\beta}}; \ x > 0$$

(i) let $x \in \mathbb{R}$.

$$F_X(x) = P(X \leq x) = \int_0^x \frac{\alpha}{\beta} y^{\alpha-1} e^{-\frac{y^\alpha}{\beta}} dy$$

$$= 1 - e^{-\frac{x^\alpha}{\beta}}; \ \forall x > 0$$

(ii) To compute F_X^{-1}: $y = 1 - e^{-\frac{x^\alpha}{\beta}} \Rightarrow x = (-\beta\log(1-y))^{\frac{1}{\alpha}}$

(iii) Let's $U \sim U(0,1)$, and set

$$Y := F_X^{-1}(U) = [-\beta\log(1-U)]^{\frac{1}{\alpha}} =_d X \sim Weinbull(\alpha, \beta)$$

$$=_d (-\beta\log(U))^{\frac{1}{\alpha}}$$

Therefore, if $u_1, \cdots u_N$ be a sample of $U(0,1)$, then

$$(-\beta\log(u_1))^{\frac{1}{\alpha}}, \cdots, (-\beta\log(u_N))^{\frac{1}{\alpha}}$$

is a sample for weinbull (α, β)

Ex: [Discrete Distribution]

$$X = \begin{cases} x_1 & \text{with probability } p_1 \\ x_2 & \text{with probability } p_2 \\ x_3 & \text{with probability } p_3. \end{cases} \qquad p_1+p_2+p_3=1$$

Let's assume $x_1 < x_2 < x_3$.

(i) $$F_X(x) = \begin{cases} 0, & \text{if } x < x_1 \\ p_1 & x_1 \le x < x_2 \\ p_1+p_2 & x_2 \le x < x_3 \\ 1 & x \ge x_3 \end{cases}$$

(ii) $$F_X^{-1}(y) = \begin{cases} x_1 & y \le p_1 \\ x_2 & p_1 < y \le p_1+p_2 \\ x_3 & y > p_1+p_2 \end{cases}$$

The generalized inverse of F_X

(iii) Take $U \sim U(0,1)$

$$Y := F_X^{-1}(U) = \begin{cases} x_1 & \text{if } U \le p_1 \\ x_2 & \text{if } p_1 < U \le p_1+p_2 \\ x_3 & \text{if} \end{cases}$$

$$\stackrel{d}{=} X.$$

Ex: [Discrete Uniform Distribution]

Let $\{x_1, \cdots x_N\}$ be a finite number of N distinct points, we want to generate a uniform r.v X on the set $\{x_1, \cdots x_N\}$, that is $1 \le k \le N$ $X = x_k$. with probability $\frac{1}{N}$, by using a uniform $U(0,1)$ distribution.

we define $X = x_k$ if $k = \lfloor NU+1 \rfloor$ for $1 \le k \le N$ (*)

Next, we check that X follows a discrete uniform distribution on $\{x_1, \cdots x_N\}$.

$$\begin{aligned} P(X = x_k) &= P(\lfloor NU+1 \rfloor = k). \\ &= P(k \le NU+1 < k+1) \\ &= P\left(\tfrac{k-1}{N} \le U < \tfrac{k}{N}\right) \\ &= \left(\tfrac{k}{N}\right) - \left(\tfrac{k-1}{N}\right) = \tfrac{1}{N}; \quad 1 \le k < N \end{aligned}$$

Remark: [How to generate a uniform sample?]

We always assume that there is a machine that produces for us a sample of uniform (a,b) for every given $(a,b) \subseteq \mathbb{R}$, that $\{u_i : i \ge 1\}$ means for every given subinterval $(c,d) \subseteq (a,b)$, it holds

$$\frac{\#\{u_i : c \le u_i \le d;\ i \le n\}}{n} \underset{\text{for large } n}{\approx} \frac{d-c}{b-a}.$$

"random" one has to be careful

Reference: Tezuka, S. (1995)
Uniform Random Numbers: Theory and practice.

Kluwer Academic Publisher Boston

Example: MATLAB call rand(...).

basic Idea: find an alternative probability distribution G, with density function g is "close" to f.

Therefore, we introduce acceptance-rejection method. Let us assume we want to generate a random sample for a r.v. whose PDF is f. Let g be another PDF, which one can properly simulate a sample from, and assume it is close to f in the following sense

Fx.

$$\exists c > 0, \frac{f(x)}{g(x)} \leq c$$

$$\sup_x \frac{f(x)}{g(x)} \leq c.$$

The acceptance-rejection method is as follows,

Algorithm 8 Acceptance rejection algorithm.

① Generate a sample y_1 of G.

(1) Generate a sample $y_1, ..., y_n$ from g.

(2) Generate a uniform $U(0,1)$ sample, $u_1, ..., u_n$ independent from $y_1, ..., y_n$.

(3) For any $1 \leq i \leq n$, if $u_i \leq \frac{f(y_i)}{cg(y_i)}$ then accept y_i as a sample of X; otherwise reject.

---

repeat.

Remark: The number of calls (denoted by N) at step 1) & 2). to successfully generate a sample for F. follows a geometric distribution with success probability $p = P(U \leq \frac{f(Y)}{cg(Y)})$ where $U \sim U(0,1)$, and $Y \sim G$.

$P(N=n) = p(1-p)^{n-1}$; $\forall n \geq 1$

Note that $E(N) = \frac{1}{p}$.

Here now we present a proof of why this algorithm has to work. First observe that

$$\begin{aligned}
P\left(U \leq \frac{f(Y)}{cg(Y)}\right) &= \int_{-\infty}^{\infty} P\left(U \leq \frac{f(Y)}{cg(Y)} | Y = y\right) g(y) dy \\
&= \int_{-\infty}^{\infty} P\left(U \leq \frac{f(Y)}{cg(Y)} | Y = y\right) g(y) dy \\
&= \int_{-\infty}^{\infty} \frac{f(y)}{c\cancel{g(y)}} \cancel{g(y)} dy \\
&= \frac{1}{c} \int_{-\infty}^{\infty} f(y) dy \\
&= \frac{1}{c}.
\end{aligned}$$

Now let us see what a random variable Y condition to $U \leq \frac{f(Y)}{cg(Y)}$ looks like:

$$P\left(Y \leq y | U \leq \frac{f(Y)}{cg(Y)}\right) = \frac{P\left(Y \leq y; U \leq \frac{f(Y)}{cg(Y)}\right)}{P\left(U \leq \frac{f(Y)}{cg(Y)}\right)}$$

Summary

① The Inversion Method

$\to F^{-1}(U) =_d F$ → both of them are discrete or continuous.

Increasing function prob. dist

$U \sim U(0,1)$ uniform

缺点: need to have connection for F inverse. 现实中很少有.

Example: $X \sim N(\mu, \sigma^2)$

① $X \overset{d}{=} \mu + \sigma Z$, where $Z \sim N(0,1)$.
the same

any normal dist 可转化为 standard dist

$F_Z(x) = P(Z \le x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{y^2}{2}} dy.$ → no way to find inverse for these.

To find the inverse: not possible!

只能 price we need to pay for any point discrete.
⇒ very expensive.
⇒ 选择别的模型

It's not simple to analytically compute F_Z^{-1}, but for every given $y \in (0,1)$,
we can approximate by using some root finding algorithm) $F_Z^{-1}(y)$.
However this approach is very expensive computationally.

To overcome: Use the 2nd Method:

② Acceptance-Rejection Method

Assume, we are interested in sampling

$X \sim F \sim f$ → continuous. → p.d.f.

Try to sample X to be continuous.

We are going to consider an alternative distribution

not far from each other. → $Y \sim G \sim g$, and we know already how to generate a sample out of G, and F and G are "close" to each other in the sense that

$$\sup_x h(x) = \sup_x \frac{f(x)}{g(x)} \le c.$$ → 因为有 sup

p.d.f doesn't hit 0.

Then, the acceptance-rejection algorithm is as follows:

① Generate a sample y of G. → alternative
② Generate sample u of $U(0)$. → uniform
} indpen.

③ If $u \le \frac{f(y)}{cg(y)}$, they accept y as a sample for F, otherwise, back to ①

How many time need to work out? ⇒ 预知

⇒ 预知 dist. ⇒ geometric dist (until first success is coming

Remark: (a) The number of call N at step ① and step ② to successfully generate a sample of dis F is a geometric distribution { parameter ? probability of success. p

$P = P\left(U \le \frac{f(Y)}{cg(Y)}\right)$, where $U \sim U(0,1)$ and $Y \sim G$.

It means that

$\forall n \ge 1 \quad P(N=n) = p(1-p)^{n-1}.$

Important: average

$E(N) = \frac{1}{p}$ (check it!)

$p = P\left(U \le \frac{f(Y)}{cg(Y)}\right)$ To conditional Y and then to unconditional to Y.

$= \int_{\mathbb{R}} P\left(U \le \frac{f(Y)}{cg(Y)} \middle| Y = y\right) g(y)\, dy.$

$P\left(U \le \frac{f(y)}{cg(y)}\right) = \frac{f(y)}{cg(y)}.$

uniform dist. less than a number is the number itself.

$= \int_{\mathbb{R}} \frac{f(y)}{cg(y)} g(y)\, dy = \frac{1}{c}.$ → density function.

$\Rightarrow p = c^{-1} \le 1$
因为是 probability
$\Rightarrow c \ge 1.$

Proof: (why A-R Algorithm work)

We need to prove that

$P(\underbrace{Y \le y}_{A} \mid \underbrace{U \le \frac{f(Y)}{cg(Y)}}_{B}) = P(X \le Y)$

accept a sample out of Y for X as same as the example going to happen X~Y.

$P(A|B) = P(B|A)\frac{P(A)}{P(B)}$ ⇒ count. dist.

$P(U \le \frac{f(Y)}{cg(Y)} \mid Y \le y) = \frac{P(U \le \frac{f(Y)}{cg(Y)}, Y \le y)}{P(Y \le y)}$

$= \frac{1}{G(y)} \int_{-\infty}^{y} \underbrace{P(U \le \frac{f(Y)}{cg(Y)} \mid Y = w)}\, \cancel{g(w)}\, dw$

$G(y) \rightarrow P(Y \le y)$

$= P(U \le \frac{f(w)}{cg(w)}) = \frac{f(w)}{c\cancel{g(w)}}$

$= \frac{1}{cG(y)} \int_{-\infty}^{y} f(w)dw$

$= \frac{F(y)}{cG(y)}$

proof is completed by noting that

$P(B) = P(U \le \frac{f(Y)}{cg(Y)}) = \frac{1}{c}$

$\frac{1}{c} = P(B)$

$G(Y) = P(Y \le y)$

$\therefore \frac{F(y)}{cG(y)} \times cG(Y)$

$= F(y)$.

最好的 c：$c \to 1 \Rightarrow P \to 1$

则 $G \to f$ 但不可能.

所以 c 越接近 1 越好. $c \approx 1$!

Example: Use acceptance-rejection Method to generate a sample for $Z \sim N(0,1)$

Since Z is a symmetric distribution, meaning that

$\underbrace{P(Z \in A) = P(Z \in -A)}_{\text{shape of density}}, \forall A \subseteq \mathbb{R}$. 因为 standard normal distribution: symmetric normal distribution → To absolute vale.

This implies that $Z =_d \underbrace{\varepsilon}_{} \underbrace{|Z|}_{\text{positive}}$, where $\varepsilon \underset{\text{independent}}{\perp} Z$, and $\varepsilon = \pm 1$ w.p. $\frac{1}{2}$

check it!

Q: how to generate absolute vale of Z?

a sample for $|Z| \ge 0$ a.s. (Amost surely.)

Note that $f_{|Z|}(x) = \frac{2}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$; $x \ge 0$.

we are going to take exponentional distribution $G \sim Exp(\lambda = 1)$, and

$g(x) = e^{-x}$; $x \ge 0$.

as alternative distribution. [Note that we know already how to generate $Exp(\lambda=1)$ by inversion method)

Next: $h(x) := \frac{f(x)}{g(x)} = \sqrt{\frac{2}{\pi}} e^{x - \frac{x^2}{2}}$; $x \ge 0$.

To max, $x=1$

→ we are looking a positive random objective ⇒ the most common one is exponential using one uniform → know how to generate that.

So, $c = \sqrt{\frac{2e}{\pi}} \approx 1.32$.

Hence, the acceptance probability is $p = c^{-1} \approx 76\%$.

Next, $\frac{f_{|Z|}(y)}{c\, g(y)} = e^{-(y-1)^2/2}$. The Algorithm is.

① Generate $Y \sim Exp(\lambda=1)$, it means that generate $U_1 \sim U(0,1)$ and set

$Y = -\log(U_1)$ ☆.

② Generate $U_2 \sim U(0,1)$

③ If $U_2 \le e^{-(Y-1)^2/2}$, set $|Z| = Y$, otherwise back to step ① → obtain a sample of absoluted vale of standard Gussian, not the standard Gussian itself.

④ Generate $U_3 \sim U(0,1)$, set $Z = |Z|$ if $U_3 \le \frac{1}{2}$ (Bernoulli)

and $Z = -|Z|$ if $U_3 > \frac{1}{2}$ ← move to itself.

Remark: Back to step ③

$U_2 \le e^{-(Y-1)^2/2} \xLeftrightarrow{\text{iff}} \underbrace{-\log(U_2)}_{=d\,Exp(\lambda=1)} \ge \underbrace{\frac{(Y-1)^2}{2}}_{\text{Independent.}}$

Try to simplify / optimize

Therefore, we can modify one algorithm as follows:

$(Z = d \cdot \varepsilon |Z|$. ($\varepsilon$: generate, Bernolli; $|Z|$: generate.)

Algorithm:

① Generate 2 independent $Exp(\lambda=1)$, i.e.

$Y_1 = -\log(U_1)$, and $Y_2 = -\log(U_2)$

② If $Y_2 \ge \frac{(Y_1-1)^2}{2}$, set $Y_1 = |Z|$, otherwise back to step ①

③ Generate $U_3 \sim U(0,1)$, and set

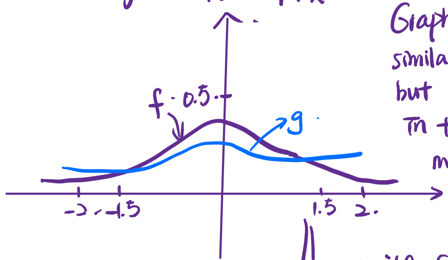
$Z = |Z|$ if $U_3 < \frac{1}{2}$.

and $Z = -|Z|$ if $U_3 > \frac{1}{2}$

This time, we are going to use a different alternative distribution g to sample standard Gaussian Distribution $Z \sim N(0,1)$, $f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

$Y \sim$ Cauchy with density function

$$g(x) = \frac{1}{\pi} \frac{1}{1+x^2}.$$



Graph of density is similar to the Gassian, but a lot of heavy tat in the tail $\Rightarrow$ variance in the moment will go to infinite. Infinite mean.

$\hookrightarrow$ quite slow to 0 compare to explanation.

$$c = \sup_{x\in\mathbb{R}} \frac{f(x)}{g(x)} = \sup_{x\in\mathbb{R}} \underbrace{\sqrt{\frac{\pi}{2}} e^{-\frac{x^2}{2}} (1+x^2)}_{\substack{\text{maximize it.} \\ \text{max at } x=\pm 1.}}.$$

$\Rightarrow$ ratio is going to bound $\to$ how good?

$\le \sqrt{\frac{\pi}{2}} e^{-\frac{1}{2}} \approx 1.52.$ pretty close to 1.

The success probability now is $p = e^{-1} \approx 66\%.$ to accept.

$76\% > 66\%$. choose experential.

To compute the acceptance-rejection algorithm, we need to understand how to generate cauchy distribution $Y \sim g$.

$\hookrightarrow$ Use the inversion method.

Let's apply the inversion Method:

$G_Y(x) = P(Y \le x) = \frac{1}{\pi} \int_{-\infty}^{x} \frac{dy}{1+y^2}. = \frac{1}{2} + \frac{1}{\pi} \arctan(x) \quad (=y)$

Then try to find $G^{-1}(y)$

$G^{-1}(y) = \tan\left(\pi\left(y - \frac{1}{2}\right)\right)$; $y \in (0,1)$.

Finally, the algorithm is as follows:

generate a sample for cauchy distribution.

① Generate a sample Y; meaning that we need to generate $U_1 \sim U(0,1)$, and set $Y = \tan(\pi(U_1 - \frac{1}{2}))$.

② Generate $U_2 \sim U(0,1)$, and if

$U_2 \le \frac{f(Y)}{cg(Y)}$, then accept Y as a sample for Z, otherwise we go back to step ①.

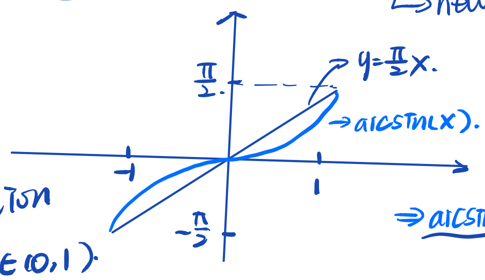
$\hookrightarrow$ standard normal

(we have 66% to accept).

Q: Can you use standard Gaussian distribution as an alternative in the acceptance-rejection method to generate Cauchy distribution?

Example: Apply the acceptance-rejection method to generate a sample for probability distribution $X \sim f$ with p.d.f

$$\underbrace{f_X(x) = \frac{2\arcsin(x)}{\pi - 2}\ ; \quad 0 < x < 1.}$$ ↳ new density function.



(*)
$\Rightarrow \arcsin(x) \leq \frac{\pi}{2}x, \quad \forall x \in (0,1).$

We define the alternative distribution

$G \sim g$ with $g(x) = 2x$; $x \in (0,1)$.

and moreover,

$$\sup_{x\in(0,1)} \frac{f(x)}{g(x)} = \sup_{x\in(0,1)} \frac{2\arcsin(x)}{(\pi-2)(2x)} \overset{\text{by } (*)}{\leq} \frac{\pi}{2(\pi-2)} \approx 1.37 = c$$

Next step is whether or not we can easily generate alternative dist. G:

$$\underbrace{G(x)}_{Y} = P(Y \leq x) = \int_0^x g(y)\,dy = \int_0^x 2y\,dy = x^2 \quad (= y).$$

$$G^{-1}(y) = \sqrt{y}.$$

Finally, the algorithm is:

step ①: Generate a sample Y; meaning generate a uniform $U_1 \sim U(0,1)$ and set $Y = \sqrt{U_1}$

step ② Generate $U_2 \sim U(0,1)$, and if $U_2 \leq \dfrac{f(\sqrt{U_1})}{c\,g(\sqrt{U_1})}$, they accept Y as a sample for X; otherwise, we go back to step ①

$$= c\int_{-\infty}^{\infty} P\left(U \leq \frac{f(Y)}{cg(Y)} | y = w \leq y\right) g(w)dw$$

$$= c\int_{-\infty}^{y} P\left(U \leq \frac{f(w)}{cg(w)}\right) g(w)dw$$

$$= c\int_{-\infty}^{y} \frac{f(w)}{c\cancel{g(w)}}\cancel{g(w)}dw$$

$$= \int_{-\infty}^{y} f(w)dw = F(y)$$

3.3.1. Combining inversion and acceptance-rejection methods. Usually, we need to combine the inversion and acceptance-rejection method, since we can generate a sample associated with PDF g by inversion method. For that, one needs to generate two independent $U(0,1)$ sample, one that by using inversion method generates a sample with PDF g, and the other one to be used in the acceptance-rejection method. So, we have the following algorithm

Algorithm 9 Inversion and acceptance-rejection, combined.

(1) Generate two $U(0,1)$ and independent random samples $u_1, ..., u_n$ and $v_1, ..., v_n$.
(2) Generate a random sample of CDF, G (PDF, g), $y_1 = G^{-1}(u_1), ..., y_n = G^{-1}(u_n)$.
(3) For any $1 \leq i \leq n$, if $v_i \leq \frac{f(y_i)}{cg(y_i)} = \frac{f\left(G^{-1}(u_i)\right)}{cg(G^{-1}(u_i))}$ then accept $y_i = G^{-1}(u_i)$ as a sample of X; otherwise reject.

EXAMPLE 48. We want to generate a sample of $|N(0,1)|$. Let z be a $N(0,1)$ random variable. We first want to generate a sample for $X = |Z|$. In that case,

$$f(x) = \frac{2}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}, x > 0.$$

Let g be $g(x) = e^{-x}$. So we have to first check whether $\frac{f(x)}{g(x)}$ is bounded or not,

$$\frac{f(x)}{g(x)} = \frac{\frac{2}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}}{e^{-x}} = \frac{2}{\sqrt{2\pi}} e^{x - \frac{x^2}{2}}.$$

Since maximum of $x - \frac{x^2}{2}$ happens at $x_0 = 1$, then,

$$\max \frac{f(x)}{g(x)} = \frac{2}{\sqrt{2\pi}} e^{(1-\frac{1}{2})} = \frac{2e^{\frac{1}{2}}}{\sqrt{2\pi}} \leq 1.4$$

So let c any thing larger than 1.4. Now from previous part we know how to generate a random sample for g. Let U, V be two independent $U(0,1)$ random variables. U will be used to generate a random sample for distribution g and V the one that is used in. Therefore, we can design the following algorithm

1-Generate two independent $U(0,1)$ random samples $u_1, ..., u_n$ and $v_1, ..., v_n$.

2-Let $y_i = \ln \frac{1}{1-u_i}$.

3-For any $1 \leq i \leq n$, if, $v_i \leq \dfrac{\frac{2}{\sqrt{2\pi}} e^{-\frac{1}{2} y_i^2}}{1.4 e^{-y_i}} = \dfrac{\frac{2}{\sqrt{2\pi}} e^{-\frac{1}{2}(\ln(1-u_i))^2}}{1.4\,(1-u_i)}$ accept $y_i = \ln \frac{1}{1-u_i}$, otherwise, reject.

Remark 49. For generating a normal distribution, we only need to have a Bernoulli random sample $(b_i)_{i=1}^n$ with values $-1, 1$ and then a normal sample is given by $(x_k b_k)_{k=1}^n$.

3.4. Variance reduction method

3.4.1. Central Limit Theorem. [CTL]

$\{X_n\}$. i.i.d random variable.

$\mu = E(X_1)$ and $\sigma^2 = Var(X)$

DEFINITION 50. Let $X_1, X_2, X_3...$ be a sequence of random variables with CDF $F_{X_1}, F_{X_2}, F_{X_3}....$ We say that $X_1, X_2, X_3...$ converge in distribution to a random variable X if, $\forall t$; $F_{X_1}(t), F_{X_2}(t), F_{X_3}(t)...$ converge to $F_X(t)$.

set $\bar{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k$, $\frac{\bar{X}_n - E(\bar{X}_n)}{\sqrt{Var(\bar{X}_n)}}$ $\to$ 0 mean. s.d. $\xrightarrow[\text{as } n\to\infty]{D} N(0,1)$.

For example, let $X_1, X_2, ...$ be a sequence of random variables with the following set of CDF:

Recall that $E(\bar{X}_n) = \mu$, and $Var(\bar{X}_n) -$

$$F_{X_n}(x) = \begin{cases} 1 - \frac{n-1}{n}\frac{1}{x} & \text{if } x \geq \frac{n-1}{n} - 1 - \frac{1}{n} \\ 0 & \text{o.w.} \end{cases}.$$

Then, we have

$$F_{X_n}(x) = 1 - \frac{(n-1)}{n}\frac{1}{x} \overset{n\to\infty}{\longrightarrow} 1 - \frac{1}{x}, x \geq 1.$$

Therefore if X has a CDF$F_X(x) = 1 - \frac{1}{x}, x \geq 1$, and 0 otherwise, then $X_n \to X$ in distribution. We show this convergence by, $X_n \xrightarrow{d} X$. The definition of converge in distribution is equivalent to

$$X_n \xrightarrow{d} X \iff E(f(X_n)) \to E(f(X)), \forall f.$$

THEOREM 51. *(Central Limit Theorem) Let $X_1, X_2, X_3...$ be a sequence of i.i.d random variables with finite mean and variance. Then, we have*

$$F_n := \frac{\bar{X}_n - \mu}{\sigma\sqrt{n}} \xrightarrow{d} N(0,1) \quad (*)$$

$\to$ continuous distribution

where $\bar{X}_n = (X_1 + ... + X_n)/n$, $\mu = E(X_i)$ and $\sigma^2 = Var(X_i)$.

$(*)$ means $\forall x \in \mathbb{R}$, it holds that $P(F_n \leq x) \xrightarrow[\text{as } n\to\infty]{\text{converge}} P(N(0,1) \leq x)$.

One of the most important implication of the CLT is that if one can reduce the variance then mean can be estimated much more efficiently.

centralize & normalize $\Rightarrow$ always normal distribution.

To see this let us again go back to our option pricing problem. Let $X_1, X_2, ..., X_n$ be an i.i.d sequence. Then we are looking for $\mu = E(X_i)$. From CLT we know,

$\to$ price of the option.

$$\frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} \to N(0,1)$$

Recall that we were interested in computing $\mu = E(X)$.
Now, let's $X_1, X_2, \dots X_n \cdots$ be an independent copies of X, and using sample mean
$\bar{X}_n := \frac{1}{n}\sum_{k=1}^{n} X_k$ $\xrightarrow{\text{converge to}} \mu$. And LLN we can give an approximation for μ (true mean).
Now CLT tells us that

$$\frac{\bar{X}_n - \mu}{\sqrt{Var(\bar{X}_n)}} = \frac{\sqrt{n}}{\sigma}(\bar{X}_n - \mu) \overset{\text{as } n \text{ is large enough.}}{\underset{(d)}{\approx}} N(0,1)$$

$\to$ as a distribution.

Therefore, $\forall z \geq 0$, we can write

$P\left(\left| \frac{\sqrt{n}}{\sigma}(\bar{X}_n - \mu) \right| \geq z \right) \approx P(|Z| \geq z)$ $\quad Z \sim N(0,1)$.

$\Rightarrow P\left(|\bar{X}_n - \mu| \geq z \frac{\sigma}{\sqrt{n}} \right) \approx 2P(Z \geq z)$

true mean.

since Z is a symmetric distribution.

$$\Longrightarrow \frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} \approx y \sim N(0, 1)$$

Therefore, for any $\alpha > 0$, (no matter how small), letting $z_{\frac{\alpha}{2}}$ such that $P(Z > z_{\frac{\alpha}{2}}) = \frac{\alpha}{2}$.

Hence, with $z = z_{\frac{\alpha}{2}}$ we can write $P(|\bar{X}_n - \mu| > z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}) \approx \alpha$. for large n.

$$\Longrightarrow \bar{X}_n - \mu \approx \frac{\sigma}{\sqrt{n}} y \sim N\left(0, \frac{\sigma^2}{n}\right).$$

free parameter.

This tell us that $\mu \in [\bar{X}_n - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}},$

So if δ is small then with probability $1 - \delta$ $(\Phi(z_0) \equiv 1 - \delta)$

$$V \in \left(\bar{X}_n - z_{\delta/2}\frac{\sigma}{\sqrt{n}}, \bar{X}_n + z_{\delta/2}\frac{\sigma}{\sqrt{n}}\right),$$

, with probability $(1-\alpha)$

var of sample dist $\Rightarrow$ smaller $\Rightarrow$ smaller interval.

getting better and better approximation.

implying that if σ is smaller then with the same sample we have more accurate answer. Therefore if we can reduce the variance then the accuracy will increase.

Conclusion: smaller variance results in better approximation.

3.4.2. Variance reduction and control variable. As we have seen, reducing the variance of the sample helps us to more efficiently find the value of the real mean with higher certainty, by using the empirical average. The variance reduction method helps us to do that by using samples of another sample whose exact mean is known, and is highly correlated with our sample of interest. For example, if we want to estimate the mean $e^{-tT}(S - K)_+$ (which is the price of a call option), we can use a sample of $e^{-rT}S$, whose mean is known and is equal to S_0. Note that they are also highly correlated.

Now, let us assume by using the empirical average, we want to approximate the mean of a random variable Y. We want to use a sample from another random variable X whose mean is known and is highly correlated with Y. First we introduce a new random variable $Y(b)$ as $Y(b) = Y - b(X - E(X))$. Observe that

$$E(Y(b)) = E(Y) - bE(X - E(X)) = E(Y) - b(E(X) - E(X)) = E(Y).$$

Therefore $Y(b)$ has the same mean as Y and this means instead of finding the mean of Y we can find mean of $Y(b)$. On the other hand , we have

$$\begin{aligned}
Var(Y(b)) &= Var(Y - b(X - E(X))) \\
&= E(Y - E(Y) - b(X - E(X)))^2 \\
&= E(Y - E(Y))^2 - b^2E(X - E(X))^2 - 2bE(Y - E(Y))(X - E(X)) \\
&= Var(Y) + b^2Var(X) - 2bCov(X, Y).
\end{aligned}$$

Remark: one can as the sample mean, give an approximation of the variance δ^2. by using the sample variance S_n^2, defined by

$$S_n^2 := \frac{1}{n-1}\sum_{k=1}^{n}(X_k-\overline{X}_n)^2.$$

Ex: show that $E(S_n^2)=\delta^2$, and moreover using LLN to show

$$S_n^2 \xrightarrow{as\ n\to\infty} \delta^2. \quad \text{almost surely (a.s.)}$$

Next, one of the method to reduce variance of a given sample is via the method called "control variables"

Control Variable:

Remember that $\mu=E(X)$ by using generating a sample for X and use sample mean as approximating μ.

Now, let's assume there is another variable denoted by C (called control variable) so that

The minimum of $Var(Y(b))$ is attained at,

$$b^* = \frac{Cov(X,Y)}{Var(X)}$$

and after implementing one,

$$\begin{aligned} Var\left(Y\left(b^*\right)\right) &= \left(1 - \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}\right) Var(Y) \\ &= \left(1 - \rho_{X,Y}\right) Var(Y). \end{aligned}$$

Since X and Y are highly correlated then $1 - \rho_{X,Y} \approx 0$ which shows that the sample mean of $Y(b)$, which has the same mean as Y, can be found with high certainty because of lower variance. But since in most cases $Cov(X,Y)$ is not known we use a sample co-variance to find b,

$$b_n = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(X_i - \bar{X})^2}.$$

and use this for sample n. Here $\bar{x}$ and $\bar{y}$ are the empirical averages.

A variance reduction algorithm for the European call. Note that a call option is first, highly correlated with the asset prices and second, its exact mean in the Black-Schole model is exactly know as its dynamic is known:

$$\begin{aligned} E\left(S_T\right) &= E\left(S_0 e^{(r-\frac{1}{2}\sigma^2)T + \sigma W_T}\right) \\ &= S_0 e^{(r-\frac{1}{2}\sigma^2)T} E\left(e^{\sigma\sqrt{T}N(0,1)}\right) \\ &= S_0 e^{(r-\frac{1}{2}\sigma^2)T} e^{\frac{1}{2}\sigma^2} = S_0 e^{rT}. \end{aligned}$$

Algorithm 10 Using variance reduction method to price European call options.

(1) Find n sample $w_1, ..., w_n \sim N(0,1)$ (either by inversion or acceptance rejection)
(2) Let $S_{i,T} = S_0 e^{(r-\frac{1}{2}\sigma^2)T+\sigma\sqrt{T}w_i}$
(3) Let $y_i = e^{-rT}(S_{i,T} - K)_+$ and $x_i = e^{-rT}S_{i,T}$
(4) Find $b_n = \frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sum(x_i-\bar{x})^2}$
(5) Let $y_i(b_n) = y_i - b(x_i - S_0)$, where we know the mean $E(x_i) = S_0$.
(6) Find $\bar{y}_i(b_n) = \frac{1}{n}\sum y_i(b)$

Example 52. In this example, using a MTLAB code we find the price of a call option by finding the sample average directly and by the algorithm above. For $S_0 = 1$, $\sigma = 0.3$, $r = 0.045$, $T = 2$, we generate a $N(0,1)$ sample of 100,000. The results of the prices from the two methods are 0.2062 and 0.2076 respectively, and the sample variances are 0.1413 and 0.0155 respectively.

3.5. Integral by Monte Carlo methods

Assume we want to find the area below a non negative function $f : [a, b] \to R$. One can use Monte-Carlo Methods to do this. The idea is to find the fraction of point of a two-dimensional uniform sample which lies below f. Let $M \geq f(x)\ \forall x \in [a, b]$. Let $U \sim U(0, M)$, $V \sim U(a, b)$ be independent, Then,

$$
\begin{aligned}
P(U \leq f(V)) &= \int_a^b P(U \leq f(v)|V = v)P(V = v)dv \\
&= \int_a^b P(U \leq f(v))\frac{1}{b-a}dv \\
&= \int_a^b \frac{f(v)}{M}\frac{1}{b-a}dv \\
&= \frac{1}{M(b-a)}\int_a^b f(v)dv
\end{aligned}
$$

$$\int_a^b f(v)dv = M(b-a)P(U \leq f(v)).$$

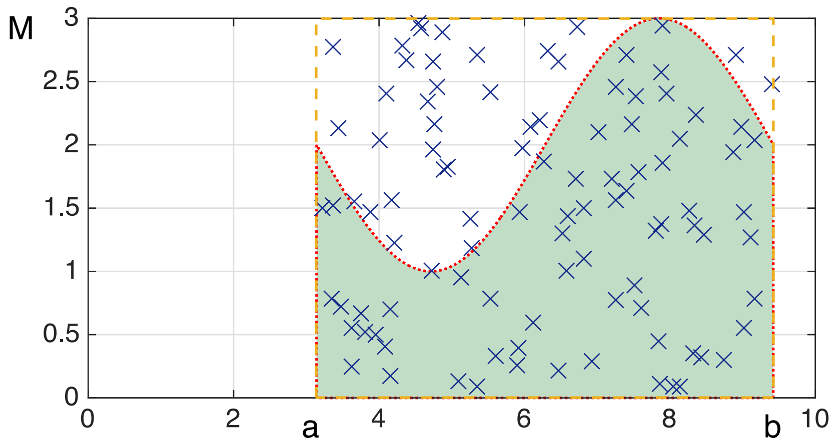


FIGURE 3.5.1. Finding integral of a function using Monte-Carlo method.

Algorithm 11 Algorithm for finding integral of a function using Monte-Carlo method.

(1) Generate $(u_i)_1^n \sim U(0, M)$, $(v_i)_1^n \sim U(a, b)$ independently.
(2) Compute $a_n = \frac{\#\{u_i \leq f(v_i)\}}{n}$
(3) Let $A_n \approx M(b-a)a_n$

EXAMPLE 53. Let us use this method to find the value of the standard normal cumulative distribution function $\Phi(1)$. For that we have to find $\frac{1}{2} + \frac{1}{\sqrt{2\pi}} \int_0^1 e^{-\frac{x^2}{2}} dx$. By setting $M = 2$, after generating a sample of the size $10^8$for u and v, one gets 0.841340640000000. The difference between this value and the real value of $\Phi(1)$ is 4.106068542952812e-06. This means for accuracy of 10^6, one need to make samples larger than 10^8.

3.6. Exercises

Exercise 54. Let $(u_i)_{i=1}^N$ be samples from $U(0,1)$.

1) Show that if for some distribution function $F:\mathbb{R}\to[0,1]$ we have the inverse F^{-1} then $(x_i)_{i=1}^N$ given by $x_i = F^{-1}(u_i)$ are distributed according to the distribution F.

2) Let X be a continuous random variable with the density

$$f(x) = \frac{1}{2}\exp(-|x|), \quad x \in \mathbb{R}.$$

Calculate the inverse of the distribution function F of the random variable X. Say that we have generated random samples $u_1 = 1/2$, $u_2 = 1/4$ and $u_3 = 3/4$ from the uniform distribution. Calculate x_1, x_2 and x_3 such that they are samples from the distribution of the random variable X.

Exercise 55. Let X be a random variable whose CDF, F_X, is continuous, strictly increasing and assume f_X is symetric. Let $r > 0$ and $b > 0$ be two fixed numbers.

1) Let $Y = |X|^{\frac{1}{r}}$. Find F_Y and its inverse F_Y^{-1}, in terms of F_X and F_X^{-1}, respectively.

2) Denoting the X's PDF by f_X, if we assume $f_X(x) = \frac{b}{2}e^{-b|x|}$, find F_X^{-1} and F_Y^{-1}.

3) Considering a $U(0,1)$ sample $u_1 = 1/4, u_2 = 1/2$ and $u_3 = 3/4$, generate a sample for Y and compute the empirical mean of Y.

Exercise 56. Let $H : \mathbb{R} \to \mathbb{R}_+$ be a strictly increasing and positive function such that $\int_{-\infty}^{\infty} H(-|x|)dx = 1$. Consider that we can generate easily a random sample of a random variable whose PDF is equal to $g(x) = H(-|x|)$. By using acceptance-rejection method, we want to generate a random sample of a random variable whose PDF is $f(x) = \frac{1}{c}H\left(-|x| - \frac{x^2}{2}\right)$, where $c = \int_{-\infty}^{\infty} H\left(-|x| - \frac{x^2}{2}\right)dx$.

(1) Show that $\frac{f(x)}{g(x)} \leq \frac{1}{c}$.

(2) Assume $H(x) = \frac{1}{2}e^x$. First, find $c = \int_{-\infty}^{\infty} H\left(-|x| - \frac{x^2}{2}\right) dx$. Then by using exercise 55, introduce an acceptance-rejection method to generate random sample with PDF, g.(The CDF of a $N(0,1)$, can be denoted by Φ)

Exercise 57. The beta distribution is a family of continuous probability distributions defined on the interval $[0,1]$ parametrized by two parameters n,$m \in \{0,1,2,3,...\}$, whose probability density function is given as follows

$$f_{n,m}(x) = b_{n,m}x^n(1-x)^m,$$

where $b_{n,m} = \left(\int_0^1 x^n(1-x)^m dx\right)^{-1}$. We denote this distribution with $\beta(n,m)$.

a) Find the cumulative distribution function $F_{n,0}$ of a $\beta(n,0)$ random variable. Note that you also have to find $b_{n,0}$ in this part.

b) Consider a $U(0,1)$ sample $u_1,...,u_k$. Using inversion method and the results in part 1), find a sample for $\beta(n,0)$ distribution.

c) Now we want to use the acceptance-rejection method to generate a sample for a general $\beta(n,m)$ distribution. For that first find the smallest number c such that $\forall x, \frac{f(x)}{g(x)} \leq c$, where $f = f_{n,m}$ and $g = f_{n,0}$. Then design an acceptance-rejection algorithm to generate a sample for $\beta(n,m)$.(in this part assume that you know $b_{n,m}$).

d) Two independent $U(0,1)$ samples have been generated. The results are $u_1 = 0.01, u_2 = 0.04, u_3 = 0.81$ and $v_1 = 0.64, v_2 = 0.49, v_3 = 0.04$. Use these samples and the results of part 2) to generate a sample for $\beta(1,2)$.

EXERCISE 58. Let us assume that X is a non-negative random variable, whose PDF is given as

$$g(x) = L\frac{x^k}{1+x^{2k+2}}, x \geq 0 \text{ and } 0 \text{ otherwise,}$$

where $L = \frac{1}{\left(\int_0^\infty \frac{x^k}{1+x^{2k+2}}\right)}$.

a) If we denote the CDF of X by G_X and its inverse by G_X^{-1}, find G_X^{-1}, and explain how one can generate random sample with CDF equal to G_X if we are given a $U(0,1)$ sample $u_1, ..., u_n$.

b) Let us assume there is another PDF described as follows

$$f(x) = \frac{ax^k}{P_{2k+1}(x) + x^{2k+2}}, x \geq 0$$

where $P_{2k+1} \geq 1$ is a non-negative function so that $P_{2k+1}(x)/x^{2k+1}$ is bounded.

Let $c = \min\{b > 0 | \forall x \geq 0, f(x)/g(x) \leq b\}$. Show $c = a/L$.

c) If you can generate independent $U(0,1)$ samples, using the acceptance-rejection method for $f(x)/cg(x)$, design an algorithm which generates random samples with PDF, f.

d) Now consider the following cases we generate two independent $U(0,1)$ sample $u_1 = 1/3$, $u_2 = 5/12$, $u_3 = 1/2$, $u_4 = 7/12$, $u_5 = 2/3$ and $.v_1 = v_3 = v_5 = 0$, $v_2 = v_4 = 1$.Generate a random sample with PDF g and a sample with PDF f for the following cases

(1) If $k = 0$ and $f(x) = \frac{a}{\sqrt{1+x}+x^2}$.
(2) If $k = 1$ and $f(x) = \frac{ax}{\sqrt{1+x^3+x^4}}$.

EXERCISE 59. A function $f : \mathbb{R} \to \mathbb{R}$ is even if $f(x) = f(-x), \forall x \in \mathbb{R}$ and is odd if $f(x) = -f(-x), \forall x \in \mathbb{R}$. Let us consider a random variable X whose PDF is an even function i.e. $f_X(x) = f_X(-x)$.

a) Show if h is an odd function then $E(h(X)) = 0$.

b) Consider an even function g. Let $Y = g(X)$. We want to use X to estimate $E(Y)$ in a Monte-Carlo simulation framework. Consider the control variable $Y_b = Y - b(X - E(X))$. Show that there is no benefit of using Y_b to find $E(Y)$ (i.e., there is no variance reduction).

c) Now consider g is odd. Find the optimal b^* that minimizes $b \mapsto \sigma_b$, where σ_b denotes the standard deviation of Y_b.

d) If in part c) we further consider $g(|x|) \leq \frac{1}{2}|x|$, show that $b^* \leq \frac{1}{2}$.

Exercise 60. Let us denote the return of a stock with X. The Probability Distribution Function of X, denoted by f, is unknown. However, according to some expert opinion it is a symmetric function around 0 i.e., $f(x) = f(-x)$. Now the company is interested to find the average positive returns. This means it is looking for $E(Y)$, when $Y = X1_{X\geq 0}$. Now we want to use X as a control variable to find the mean of Y. Let us assume $(X_i, Y_i), i = 1, ..., N$ are i.i.d copies of (X, Y). Let us introduce a new random variable $Y(b)$, for any positive number b as follows

$$Y(b) = Y - bX.$$

Answer the following questions regarding $Y(b)$

a) Show $E(X) = 0, i = 1, .., N$ and use it to show $E(Y(b)) = E(Y)$.

b) Find the standard deviation $\sigma(Y(b))$, in terms of standard deviations $\sigma(Y)$, $\sigma(X)$ and number b.

c) Find all numbers b for which $\sigma(Y(b)) \leq \sigma(Y)$. Find b^*, the number that minimizes $\sigma(Y(b))$.

d) Express $Y(b^*)$ in terms of the random variable X. If we have a sample $X_1, ..., X_n$ from X that the first k are nonnegative and the rest are negative, what is the sample mean of Y and $Y(b^*)$.

Exercise 61. In this question we will work with Black–Scholes model. We wish to use control variable to improve Monte Carlo estimate of price of an European call option (ignoring the fact that we could simply use Black–Scholes formula). Explain, in detail, how $e^{-r(T-t)}S_T$ can be used as a control variate in the Monte Carlo estimate.

Exercise 62. Using MATLAB to apply the Monte Carlo method we discussed in the lecture to approximate the following integral:

$$\int_{-10}^{10} \frac{1}{1+0.1x^2}dx.$$

Write a computer code in MATLAB for this problem.

Our primary goal is to the numerical approximation of the (random) dynamic of stock price movement in the continuous model of trading. 动态股票连续模型. For example, as we have seen before, the dynamic of the stock price in the celebrated Black-Scholes market model is given by the following stochastic differential equation (have after SDE):

$$dS_t = \mu S_t dt + \sigma S_t dW_t; \quad t \in [0, t],$$

where

$S = (S_t)_{t \in [0,T]}$ denote the stock price movement which is a random movement.

μ: drift parameter.

σ: volatility parameter

$W = (W_t)_{t \in [0,T]}$ which is a random (stochastic) process, and is called Brownian motion.

At this moment, it is enough to know that for every given $t \in [0, T]$, the random variable $W_t \sim N(0, t)$ is distributed as a Gaussian distribution with zero mean and variance t.

Before dig into stochastic differential equations, we first treat the ordinary deterministic counterpart.
不可逆转的

CHAPTER 4

Numerical methods for ordinary and stochastic differential equations

下面红蓝部分: an initial value problem.

4.1. Ordinary differential equation (ODE)

An ordinary differential equation (ODE) of order one is defined as,

consider a first order differential equation of the form

$$(*)\quad \begin{cases} x' = f(t,x), t \in [t_0, t_M) \\ x(t_0) = x_0 \end{cases}, \quad ☆.$$

general function. T.d. variable. (time)

where $x' = \frac{dx}{dt}$, f is a function of two arguments t and x, usually smooth enough in both arguments. In this equation x is a function of t, $x(t)$. We always have to assume that an initial value for function x at a specific point t_0, $x(t_0) = x_0$ and we are concerned with an answer x which holds true on an interval $[t_0, t_M)$ where $t_0 < t_M \le +\infty$. Two main and primary theoretical issues are

general solution of (*): $x = \int f(t,x)\,dt + C$

initial condition

① $0 \le$ general time domain.

② The name "first-order" obviously comes from the fact that equation (*) only involves the first derivative of the function. $t \to x(t)$.

③ 可称 equations of type (*) ordinary linear because the apperance of derivative x' in there is linear e.g.: $(x')^2 = f(t,x)$ in case of order 1, but is not a linear equation.

2 fundamental questions to be addressed.

1) The existence of a solution.
2) The uniqueness of the solution.

Example 63. Consider the following ODE,

$$x' = ax, \;\; x(0) = x_0 > 0.$$

This problem has a unique solution, $x = x_0 e^{at}$ for $t \ge 0$. So this problem has a solution. One can also this solution is unique.

Example 64. Consider the following ODE,

$$x' = x^2, x(0) = x_0 > 0.$$

To solve this problem we have,

$$x' = x^2 \implies \frac{x'}{x^2} = 1 \implies \frac{d}{dt}\left(x^{-1}\right) = -1$$

$x(0) = \frac{1}{4}(t_0 - c)^2$

$= \frac{1}{4}c^2$

$c^2 = 4x_0$

$c = 2x_0^{\frac{1}{2}}$

$$\implies x^{-1} = -t + c$$

$$\implies x = \frac{1}{c - t}$$

$$x(0) = x_0 \implies \frac{1}{c - 0} = x_0 \implies c = \frac{1}{x_0}$$

$$\implies x = \frac{1}{\frac{1}{x_0} - t}$$

As one can see at, $t = \frac{1}{x_0}$ this problem blows-up therefore this problem has a solution on, (this solution is only well-defined if $t \neq x_0^{-1}$.

$$\left[0, \frac{1}{x_0}\right). \text{ or } [0, \tfrac{1}{x_0}) \cup (\tfrac{1}{x_0}, +\infty)$$

Example: Let $\alpha \in (0,1)$ be a fixed number. consider the initial value problem

$$\begin{cases} x' = |x|^\alpha \\ x(0) = 0 \end{cases}$$

sol: It's simple (check!) to verify that for every non-negative real number c:

$$x_c(t) = \begin{cases} (1-\alpha)^{\frac{1}{1-\alpha}}(t-c)^{\frac{1}{1-\alpha}}, & 0 \le t < \infty \\ 0, & 0 \le t \le c. \end{cases}$$

(solution to the initial value problem on the interval $[0, \infty)$)

Example 65. Consider the following ODE,

$$x' = |x|^{\frac{1}{2}}, x_0 = 0.$$

$x(t) = \frac{1}{4}(x-c)^2$.

shows existence of solution, not uniqueness.

This problem has infinite number of solutions,

$$x_c(x) = \begin{cases} \frac{1}{4}(t-c)^2 & \text{if} \quad c \le t < \infty \\ 0 & \text{with} \quad 0 \le t \le c \end{cases},$$

for any $c \in R_+$. So there is no unique solution.

[However, note that as soon as $\alpha \ge 1$, then there is a unique solution $x(t) \equiv 0$]

Theorem 66. *(Picard) Assume the following assumption hold,*
1) $\exists K, |f(t, x_0)| \le K, \forall t \in [t_0, t_M)$ suppose that the real-valued function $(t,x) \mapsto f(t,x)$ is continuous on the rectangular
2) $\exists L, C$ such that $C \ge \frac{K}{L}(e^{L(t_M - t_0)} - 1)$. Let $D = \{(t, x) | t_0 \le t \le t_M, |x - x_0| \le c\}$.
If $\forall (t, u), (t, v) \in D$, $|f(t, u) - f(t, v)| \le L|u - v|$. (a) $|f(t, x_0)| \le k$; $\forall t \in [t_0, t_M]$

(b) Lipschitz continuity assumption on the second variable.

Further. (* * *)

Then, there exists a unique function $x \in C^1[t_0, t_M]$ such tat $x(t_0) = x_0$ and $x' = f(t, x)$. Furthermore, $|x(t) - x_0| \le c$, $\forall t_0 \le t \le t_M$.

The idea of the proof. Recursively construct a sequence $x_0(t) = x_0$, (the constant function)

$$x_n(t) = x_0 + \int_{t_0}^{t} f(s, x_{n-1}(s))ds$$

First note that since f is continuous on D, so each function x_n is continuous on $[t_0, t_M]$

$$x(t) = \lim x_n(t) = \lim x_0 + \int_{t_0}^{t} \lim f(s, x(s))ds$$

$$x_{n+1}(t) - x_n(t) = \int_{t_0}^{t} \left(f(s, x_n(s)) - f(s, x_{n-1}(s)) \right) ds.$$

Next, we claim that

(i) $\forall n \ge 1$; $|x_n(t) - x_{n-1}(t)| \le \frac{k}{L} \frac{(L(t-t_0))^n}{n!}$; $t_0 \le t \le t_M$

(ii) $|X_k(t) - X_0| \le \frac{K}{L}\sum_{j=1}^{k} \frac{(L(t-t_0))^j}{j!}$; $t_0 \le t \le t_M$, $k=1,2,\cdots,n$.

Therefore, $(t, X_{n-1}(t)) \in D$, $(t, X_n(t)) \in D$ for every $t \in [t_0, t_M]$

Hence, $\forall t \in [t_0, t_M]$,

$|X_{n+1}(t) - X_n(t)| = \left|\int_{t_0}^t (f(s,X_n(s)) - f(s,X_{n-1}(s)))\, ds\right| \le \int_{t_0}^t |f(s,X_n(s)) - f(s,X_{n-1}(s))|ds$

$\le L\int_{t_0}^t |X_n(s) - X_{n-1}(s)|\, ds$

$\le \frac{K}{L}\frac{(L(t-t_0))^{n+1}}{(n+1)!}$ (by induction assumption)

$$\boxed{C \ge \frac{K}{L}\left(e^{L(t_M-t_0)} - 1\right)}$$

$$= x_0 + \int_{t_0}^{t} f(s, x(s))ds$$

$$\Rightarrow x'(t) = f(t, x(t)), x(t_0) = x_0.$$

One then can easily show that,

$$|x_n(t) - x_{n-1}(t)| \le \frac{K}{L}\frac{L(t-t_0)^n}{n!}$$

$$\Longrightarrow$$

$$\sum_{j=1}^{\infty}(x_j(t) - x_{j-1}(t)) \le \frac{K}{L}e^{L(t-t_0)}$$

$$x_n(t) - x_0(t) = \sum_{j=1}^{n}(x_j(t) - x_{j-1}(t))$$

Moreover, using the above estimate and (ii), we can write

$|X_{n+1}(t) - X_0| \le |X_{n+1}(t) - X_n(t)| + |X_n(t) - X_0|$

$\le \frac{K}{L}\frac{(L(t-t_0))^{n+1}}{(n+1)!} + \frac{K}{L}\sum_{j=1}^{n}\frac{(L(t-t_0))^j}{j!}$

$= \frac{K}{L}\sum_{j=1}^{n+1}\frac{(L(t-t_0))^j}{j!}$

$= \frac{K}{L}\left(e^{L(t_M-t_0)} - 1\right) < \infty.$

Therefore, the infinite series $\sum_{j=1}^{\infty}(X_j(t) - X_{j-1}(t))$ converges absolutely and uniformly for every $t \in [t_0, t_M]$.

where $x_n(t)$ is converging. Finally, let us check the uniqueness

$$x(t) - z(t) = \int_{t_0}^{t}(f(s, x(s)) - f(s, z(s)))\, ds$$

(***)

$$\Longrightarrow |x(t) - z(t)| \le L\int_{t_0}^{t}|x(s) - z(s)|ds$$

$$\le Lm(t - t_0), \text{ where, } m = \max_{t_0 \le t \le t_M}|x(t) - z(t)|$$

$$\Longrightarrow |x - z| \le m\frac{[L(t-t_0)]^k}{k!} \to 0, k \to \infty.$$

This yields that $y(t) = X(t)$ $\forall t \in [t_0, t_M]$, and we are done.

However note that $X_0 + \sum_{j=1}^{n}(X_j(t) - X_{j-1}(t)) = X_n(t)$

So, the sequence of function $(X_n)_{n\ge1}$ converges to a limit, let's say X, uniformly on interval $[t_0, t_M]$. Note that since each X_n is continuous, the limit function X is also continuous on $[t_0, t_M]$.

Next, $\forall t \in [t_0, t_M]$

$X(t) = \lim_{n\to\infty} X_{n+1}(t)$

$= X_0 + \int_{t_0}^t \lim_{n\to\infty} f(s, X_n(s))ds$

$= X_0 + \int_{t_0}^t \lim_{n\to\infty} f(s, X_n(s))ds$

$= X_0 + \int_{t_0}^t f(s, X(s))ds$ (****)

In the application of the Picard's theorem, it is necessary to choose a value of constant C so that various hypothesis are satisfied, in particular (***). However, it can be show that if $\frac{\partial f}{\partial x}$ is continuous in a neibourhood to $(t_0, X(t_0))$, the latter condition will satisfied if $(t_M - t_0)$ is sufficiently small.

EXAMPLE 67. Consider $x' = py + q$, where p,q are constant numbers. Observe that $f(t, x) = px + q$ is linear. Therefore, it is Lipschitz, since $|f(t, v) - f(t, u)| = |p||u - v|$ where $L = |p|$. For the bounded-ness for the first argument we need to observe that $f(t, x_0) = px_0 + q \Longrightarrow |f(t, x_0)| \le |p||x_0| + |q|$. Note that $|f(t, x_0)| \le |p||x_0| + |q|$ does not depend on t. Now, to find C we only need to choose C so that

To satisfied (***)

$$C \ge \frac{|px_0| + |q|}{|p|}\left(e^{|p|(t_M-t_0)} - 1\right).$$

$|f(t, x_0)| = |px_0 + q|$

$\le |p||x_0| + |q| =: K.$

Note also that we can take $K = |p||x_0| + |q|$.

$|f(t,u) - f(t,v)|$

$= |p(u-v)| = |p||u-v| \Rightarrow$ take $L = |p|$.

unique solution

hence, we get unique solution

$$\begin{cases} x' = Px + g \\ x(0) = x_0 \end{cases}$$ on $[t_0, t_M]$ for sufficiently large value chosen for C

EXAMPLE 68. Consider $x' = x^2$, $x(0) = 1$. Observe that, $f(t, x) = x^2$ $=:k$. For any interval $[0, t_M]$ we can take $K = 1$. If we pick a C, then we have,

$$|f(t,u) - f(t,v)| = |u^2 - v^2| = |u + v||u - v| \le L|u - v|.$$

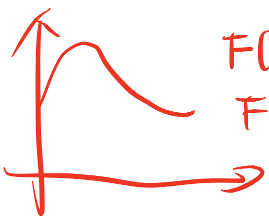
$|u|+|v| \le (C+|x_0|) + |(C+x_0)| = 2(1+C)$ $2(1+C)|u-v|$, so $L = 2(1+C)$

$$|u + v| \le |(1 + C)| + |(1 - C)| = L = 2(1 + C)$$

$$1 - C \le u, v \le 1 + C$$

$|x_0|$ $|x_0|$.

Also, we have to check,



F(1.17141) ≤ 0.46

F(C) ≤ 0.43 ∀ C ≥ 0

$$C \ge \frac{1}{2(1 + C)} \left(e^{2(1+C)t_M} - 1\right)$$

$$\implies t_M \le \frac{1}{2(1 + C)} \ln(1 + 2C + 2C^2)$$

As a result, one can not verify the existence and uniqueness of solution. But since $\frac{1}{2(1+C)} \ln(1 + 2C + 2C^2)$ is a bounded function of C, so one cannot verify if the solution is global.

$x(t) = \frac{1}{1-t}$ not continuous on any interval $[0, t_M]$ with $t_M > 1$

Picard: sufficient but not necessary for the existence and the uniqueness of the solution.

4.2. Numerical approximation scheme for ODEs

Let us divide the interval $[t_0, t_M]$ to N equi-spaced sub-intervals, or a partition, with length,

$t_n := t_0 + nh$ → mesh points

$$h = \frac{t_M - t_0}{N}.$$ → step size or mesh size.

We call $\{t_0, t_1, \ldots, t_n\}$ a grid. Grid points are given by $t_n = t_0 + nh$, and the step size is equal to h. In general, an approximation method can be written in the following general form,

→ Euler's Method.

只由前一项决定

$$x_{n+1} = x_n + h\Phi(t_n, x_n; h), \tag{4.2.1}$$, $n = 0, 1, \cdots, N-1$ (***)

$x_0 = x(t_0)$

continuous function of its variables →

where $\Phi(t, x; h)$ is a function usually based on f. The sequence $\{x_n\}_n$ is meant to be an approximation of the solution x. Note that this sequence is not necessarily equal to the values of the function; just the notation is similar. This sequence can approximate the real function x. Since the aim of a numerical method is to approximation a solution we introduce two types of errors.

In order to assess the accuracy of the numerical method (***), we define:

Error Analysis:

Global error:

$$e_n = |x(t_n) - x_n|.$$

True value of mesh point t_n.   approximate value.

Numerical Approximation schemes for ODEs

consider the following ODE (or Initial value problem)

$$(*)\begin{cases} x' = f(x,t); & t \in [t_0, t_n] \\ x(t_0) = x_0 \in \mathbb{R}. \end{cases}$$

where t is the time variable, and f is a general function.
We also assume that the assumptions of the Picard's theorem fullfill.

Our aim is to consider step-by-step numerical method for the approximate solutions of the initial value problem $(*)$

接上面的讲述

For each $n = 0, 1, \cdots, N-1$, we search a numerical approximation X_n to $X(t_n)$ the value of the analytical solution at the mesh point t_n;

these values X_n are calculated in succession for $n = 1, 2, \cdots N-1$

A one-step method express X_{n+1} in terms of the previous value X_n. [Generally, a k-step method express X_{n+1} in terms of the k previous vales $X_n, X_{n-1}, \cdots X_{n-k+1}$; for $k \geq 2$].

The simplest example of a one-step method for the numerical solution of the initial value problem. $(*)$ is Euler's Method. (4.2.1) recall: $X_{n+1} = X_n + h\, f(X_n, T_n)$

$= \Phi(t_n, X(t_n); h)$. Hence, the truncation error for Euler's method is given by

Truncation error:

$x'(t_n) = f(t_n, X(t_n))$

$$T_n = \left| \frac{x(t_{n+1}) - x(t_n)}{h} - \phi(t_n, x(t_n); h) \right| .$$

We have the following important theorem. (provide a bound on the magnitable if the global error in terms of the truncation error.

THEOREM 69. *Let $D = \{(t,x) | t_0 \le t \le t_M, |x - x_0| < c\}$ and L_ϕ be a number that for some h_0.*

$n = 1, \ldots N.$

$$\forall h \in [0, h_0], |\Phi(t, u; h) - \Phi(t, v; h)| \le L_\phi |u - v|$$

and $\forall (t,u), (t,v) \in D$. Then,

global error

$$|e_n| \le \frac{T_{max}}{L_\phi} \left(e^{L_\phi (t_n - t_0)} - 1 \right), n = 0, 1, ..., N,$$

↳ global error.

where $T_{max} = \max_{0 \le n \le N} T_n$.

$0 \le n \le N-1$ ↳ maximum of all truncation errors

Later, in few examples, we will show how we can use this theorem to find out if a problem has a solution or not.

Consistency and Convergence.

DEFINITION 70. The numerical method, $x_{n+1} = x_n + h\Phi(t_n, x_n; h)$ is consistent if,

$$\forall \epsilon, \exists h(\epsilon), 0 \le h \le h(\epsilon) \implies |T_n| \le \epsilon.$$

If we let $n \to \infty$ $h \to 0$, then, $t_n \to t$, $\lim T_n = x'(t) - \Phi(t, x(t); 0)$. Therefore $T_n \to 0$ iff $x'(t) - \Phi(t, x(t); 0)$ or $\Phi(t, x(t); 0) = f(t, x)$. This condition is equivalent to consistency.

Def: $X_{n+1} = X_n + h\Phi(t_n, X_n; h)$ is consistent with the differential equation $x' = f(t,x)$ if the truncation error T_n is such that for any $\varepsilon > 0$, $\exists$ positive $h = h(\varepsilon)$ depends on ε, for which $T_n < \varepsilon$ for every $0 < h < h(\varepsilon)$ and every pair of points $(t_n, X(t_n))$ and $(t_{n+1}, X(t_{n+1}))$

Then, according to the theorem T_n before $e_n \le \frac{T_{max}}{L_\Phi}\left(e^{L_\Phi(t_n - t_0)} - 1\right)$, T_{max} is the maximum of all truncation errors T_n; $n = 0, \ldots N-1$.

THEOREM 71. *If consistency holds then numerical solution converges to real solution,* (since assume Φ is continuous).

$\lim_{n\to\infty} x_n(t) = x(t)$, as $t_n \to t$, when $h \to 0$ or $n \to \infty$

we have $T_n \to x'(t) - \Phi(t, x(t); 0)$

$X_{n+1} = X_n + h\Phi(t_n, X_n; h)$ is constant if and only if $\Phi(t, x; 0) = f(t,x)$

⇓ consistency condition.

DEFINITION 72. The order of accuracy is the largest number p such that, $\exists K$, $T_n \le Kh^p$ Therefore for Euler explicit method we know $p \ge 1$.

$T_n \le k h^p$ for $0 < h < h_0$

Euler's method has the order ... $p = 1$.

Let $h>0$, and consider the mesh points
$t_n = t_0 + nh$; $n=0,1,\cdots N$
of the interval $[t_0, t_N]$, so h is given by
$h = \frac{t_N - t_0}{N}$
The sequence of numbers
$x_{n+1} = x_n + hf(t_n, x_n)$
$x_0 = x(t_0)$, $n = 0,1,\cdots N-1$.

of accuracy

Other one-step method: Runge-Kutta method (课后看)

4.3. Explicit method (Euler Method)

Let us consider a grid on $[t_0, t_M]$, then consider the following sequence of numbers

$$x_{n+1} = x_n + hf(t_n, x_n),$$

$\begin{cases} x' = f(t,x) \\ x(t_0) = x_0 \in \mathbb{R}. \end{cases}$

$n = 0, 1, 2, ..., N$. To see why it is an approximation, note that by the Tailor expansion if x is the real solution then,

$$x(t_n + h) = x(t_n) + hf(t_n, x(t_n)) + o(h^2).$$

Now, if we replace $x(t_n + h)$ and $x(t_n)$ by their approximations $x_{n+1}$and x_n we get the above approximation scheme.

Now, let us apply Theorem 69 for the Euler Method. Comparing the general form of a numerical scheme and the explicit method, it is clear that $\Phi(t, x; h) = f(t, x)$.

Convergence of explicit method. Now using the Taylor expansion we get

$$\exists\zeta_n \in (t_n, t_{n+1}), x(t_{n+1}) = x(t_n) + hx'(t_n) + \frac{h^2}{2}x''(\zeta_n)$$

, where $t_n < \xi_n < t_{n+1}$. we obtain that.

$$\Longrightarrow \frac{x(t_{n+1}) - x(t_n)}{h} - x'(t_n) = \frac{h}{2}x''(\zeta)$$

$$\Longrightarrow T_n = \frac{h}{2}\left|x''(\zeta_n)\right| \leq \frac{h}{2}M_2.$$

$M_2 := \max|x''(t)|$ for every $t \in [t_0, t_N]$

If $M_2 = \max_{t_0 \le \zeta \le t_M} x''(\zeta)$ then $T_n \le \frac{M_2 h}{2}$. Using in the theorem statement,

$T_{max} = \max_{0 \le n \le N-1} T_n \le \frac{h}{2}M_2.$

$e_n \le \frac{T_{max}}{L_\Phi}\left(e^{L_\Phi(t_n - t_0)} - 1\right)$

$$e_n \le \frac{1}{2}M_2\left[\frac{e^{L(t_n - t_0)} - 1}{L}\right]h.$$

$n = 0,1,\cdots N.$

$\Phi(t_n, x_n; h) = f(t_n, x_n)$

So as long as we have a bounded M_2 on an interval then the approximation by the explicit solution will converge to the real solution as $h \to 0$.

EXAMPLE 73. Consider $x' = tg^{-1}(x)$, $x(0) = x_0 \in \mathbb{R}$. To find L we use mean value theorem.

即 $\begin{cases} x' = \tan^{-1} x \\ x(0) = x_0 \in \mathbb{R}. \text{ (so, } t_0 = 0) \end{cases}$

(t, η)

$$|f(t, u) - f(t, v)| = \left|\frac{\partial f}{\partial y}(\eta)\right| |u - v|$$

$= \left|\frac{\partial f}{\partial x}(t,\eta)(u-v)\right|$

$= \left|\frac{\partial f}{\partial x}(t,\eta)\right| |u-v|.$

In order to find an upper bound on the global error
$e_n = x(t_n) - x_n$.
where x_n stands for the Euler approximation to $x(t_n)$, we need to determine the constants L and M_2 in the last estimate.

where η lies between u and v. In our case

$$\left|\frac{\partial f}{\partial x}(t,x)\right| = \left|\frac{\partial f}{\partial x}\tan^{-1}x\right| = \left|\frac{1}{1+x^2}\right| \le 1.$$

Hence, $L=1$.

Here, $f(t,x)=\tan^{-1}x$, so by the mean value theorem:

$$= \frac{1}{1+\eta^2}|u-v| \le |u-v|, \frac{1}{1+\eta^2} \le 1 = L$$

To find M_2 we need to look at the maximum of x''. From the equation we have,

$$x' = \overset{\tan^{-1}x.}{tg^{-1}x} \Longrightarrow x'' = x'\overset{\tan^{-1}x.}{\frac{1}{1+x^2}} = \frac{tg^{-1}(x)}{1+x^2}. = \frac{1}{1+x^2}\tan^{-1}x$$

Therefore, to find M_2, we have to find the maximum of $\frac{tg^{-1}(x)}{1+x^2}$. But since $|tg^{-1}(x)| \le \frac{\pi}{2}$, then an upper bound for $M_2$is $\frac{\pi}{2}$, and we get

$$M_2 = \sup_{t\in[t_0,t_M]} |x''(t)| = \sup_{t\in[t_0,t_M]} \underbrace{\left|\frac{1}{1+x^2(t)}\right|}_{\le 1} \times \underbrace{|\tan^{-1}x(t)|}_{\le\frac{\pi}{2}} = \frac{\pi}{2}$$

$$|e_n| \le \frac{1}{2}M_2(e^{t_n}-1)h \le \frac{1}{4}\pi(e^{t_n}-1)h \le \frac{1}{4}\pi(e^{t_M}-1)h.$$

Example 74. Consider the following ODE,

$$x' = 2x, \; x(0) = 1.$$

The explicit method simply gives $x_{n+1} = x_n + 2hx_n$. This problem has a unique solution, $x = e^{2t}$ for $t \ge 0$. In the following we have the graph for the explicit method for $N = 8, 16$ and 32.

Take into account that $t_0=0$, we obtain that

$$e_n \le \frac{\pi}{4}(e^{t_n}-1)h.$$

Thus, given a tolerance "TOL", specified beforhand, we can ensure that the error between the (unknown) analytical solution and it's numerical approximation does not exceed this tolerance by chosing a positive step size h, so that

$$h \le \frac{4}{\pi(e^{t_n}-1)}\,\mathrm{TOL}$$

For such h we shall have

$$|x(t_n) - x_n| = e_n \le \mathrm{TOL}$$

for $n=0,1,\dots N$

Thus, at least in principle, we can calculate the numerical solution to arbitary high accuracy by choosing a sufficiently small mesh size h.

Example: consider the IVP.

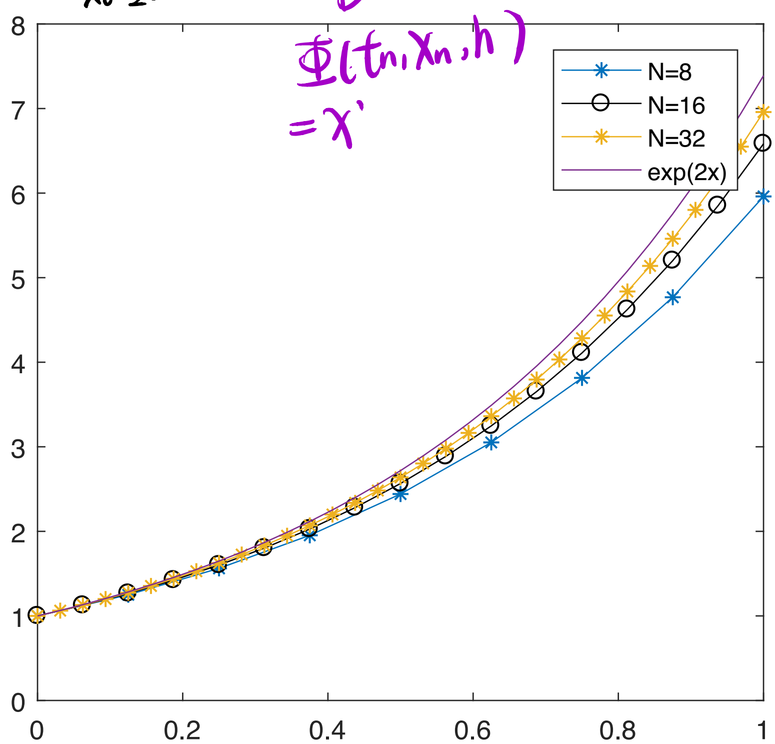
$$x' = x^2 + g(t)$$
$$x(0) = 2.$$

$g(t) = \dfrac{t^4 - 6t^3 + 12t^2 - 14t + 9}{(1+t)^2}$ is so chosen that the solution is known, and is given by

$$x(t) = \frac{(1-t)(2-t)}{1+t}.$$

Euler's method gives the approximate solution (with step size h)

$$x_{n+1} = x_n + h\underbrace{(x_n^2 + g(t_n))}_{\Phi(t_n, x_n, h) = x'}; \quad n = 0, 1, \cdots, N$$
$$x_0 = 2.$$



The results: step size $(h/2)$ gives a reduction of the error also by a factor of roughly 2, in agreement with the error bound.

$$|e_n| \le \frac{1}{2} M_2 \left(\frac{e^{L(t-t_0)} - 1}{L} \right) h$$
$$n = 0, 1, \cdots N$$

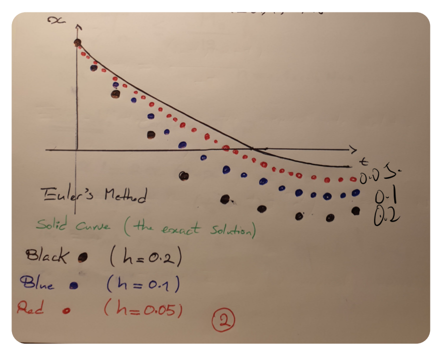


Figure 4.3.1. Explicit method.

Global error in the trapezium rule method:

$e_n \le \frac{M_3}{12 L_\Phi}\left(e^{L_\Phi(t_n-t_0)}-1\right) h^2.$

where $L_\Phi \le \frac{L_f}{1-\frac{1}{2}hL_f}$ provided that $\frac{1}{2}hL_f<1.$

4.4. Implicit method

One can find the sequence of numerical approximation as,

$$[t]\quad x_{n+1} = x_n + \frac{h}{2}\left(f\left(t_n, x_n\right)\right) + f\left(t_{n+1}, x_{n+1}\right)$$

The idea of this approximation scheme comes from Trapezoidal rule,

Trapezoidal approximation of integral:

$$x(t_{n+1}) - x(t_n) = \int_{t_n}^{t_{n+1}} x'(s)ds$$

$$\approx \frac{h}{2}\left(x'\left(t_{n+1}\right) + x'\left(t_n\right)\right)$$

$x'(t_{n+1}) = f(t_{n+1}, x(t_{n+1}))$
$\approx f(t_{n+1}, x_{n+1})$
$x'(t_n) \approx f(t_n, x_n)$

The truncation error here is then,

$$T_n = \frac{x(t_{n+1}) - x(t_n)}{h} - \frac{1}{2}[f(t_n, x(t_n)) + f(t_{n+1}, x(t_{n+1}))].$$

use h=0.4, and 0.2 larger than for Eulers method.

From trapezoidal method we know,

$$T_n \le \frac{1}{12}h^2 M_3$$

Truncation Error

The implicit method trapezium rule converges much faster than Euler method.

when $M_3 = \max x'''.$ $= \sup_{t\in[t_0,t_n]} |x'''(t)|$

The main difference between the implicit and the explicit method is that in the implicit method the term x_{n+1} appears on the right hand side. Therefore, this term need to be found by solving, usually, a non-linear root-finding problem. x_{n+1} is usually found by Newton method, with starting point $x_n + hf(t_n, x_n)$.

Example 75. Consider the following ODE $x' = 2x$, $x(0) = 1$. The implicit method simply gives $x_{n+1} = x_n + \frac{h}{2}\left(2x_n + 2x_{n+1}\right) = x_n + h\left(x_n + x_{n+1}\right)$. So, we have $x_{n+1} = \frac{1+h}{1-h}x_n$. This problem has a unique solution, $x = e^{2t}$ for $t \ge 0$. In the following we have the graph for the explicit method for $N = 8, 16$ and 32. As one can see the accuracy of this method is so high that one cannot distinguish between the numerical solutions and the real solution.

Sol: (1) unique solution is given by

$x(t) = e^{2t}.$

$\Rightarrow$ trapezium rule method

$x_{n+1} = x_n + \frac{h}{2}(2x_n + 2x_{n+1})$
$= x_n + h(x_n + x_{n+1})$
$\Rightarrow x_{n+1} = \frac{1+h}{1-h} x_n$

And Euler method:
$x_{n+1} = x_n + 2h x_n$
$= (1+2h) x_n$.

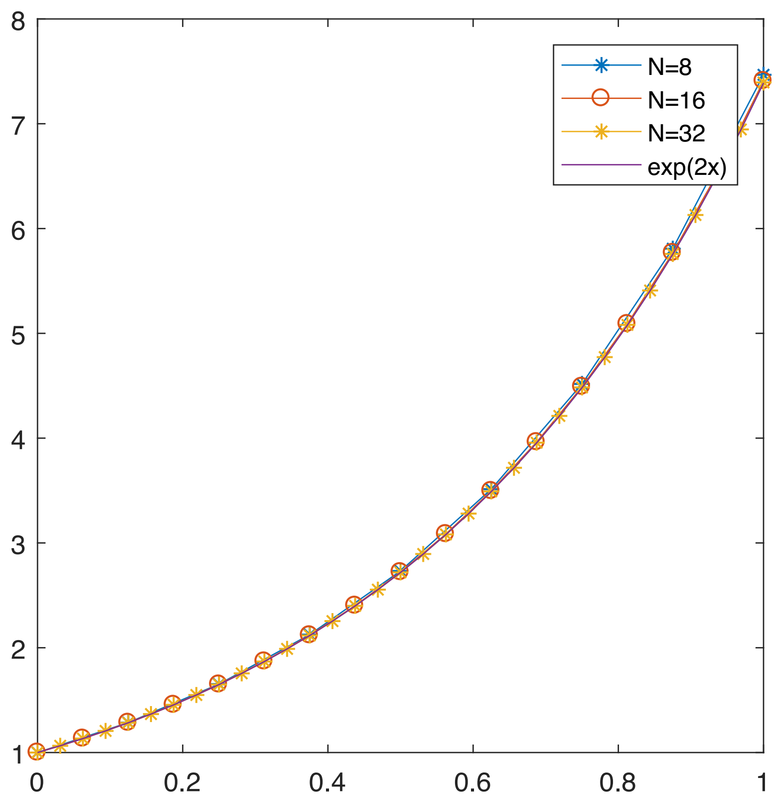


FIGURE 4.4.1. Implicit method.

Example: Consider IVP

$x' = 2x(1-x)$, $t > 10$

$x(10) = \frac{1}{5}$ $\quad (t_0 = 10)$

t_n

$x' = 2x(1-x)$

$x(10) = \frac{1}{5}$,

$(t_0 = 10)$

The Euler's Method by step size h gives the numerical solution

③

$x_{n+1} = x_n + h(2x_n(1-x_n))$

Euler's Method:

$x_{n+1} = x_n + h\Phi(t_n, x_n; h)$,

X'

$x_0 = 1/5 = 0.2$

The result with $h = 0.2$ on the time interval $t \in [10, 11]$ is as follows:

n	t_n	x_n	x_n'
0	10.0	0.2	0.32
1	10.2	0.2640	0.3886
2	10.4	0.3417	0.4499
3	10.6	0.4317	0.4907
4	10.8	0.5298	0.4982
5	11.0	0.6295	0.4665

$X(10)$ $\quad h_0$ $\quad 10+0.2$ $\quad 10+0.2+0.2$

$X_1 = 0.2 + 0.2(2 \times 0.2(1-0.2))$
$= 0.264$

$X_0' = 2 \times 0.2(1-0.2)$
$= 0.32$.

You can also compare the numerical results with the real values. To do this, you need to know that the exact solution is

$$x(t) = \frac{1}{1 + 4\exp(2(10-t))}$$

Remark:

(a) In the derivation of Euler's Method, we used the $\boxed{\text{forward}}$ difference approximation.

$$X'(t_n) \approx \frac{X(t_n+h)-X(t_n)}{h}.$$

Backward difference approximation

$$X'(t_n) \approx \frac{X(t_n)-X(t_n-h)}{h}.$$

differential equation:

$X'(t) = f(t, X(t))$ at $t=t_n$ is discretized as

$$X_n = X_{n-1} + hf(t_n, X_n)$$

shifting the index by 1, we than obtain the so-call

Backward Euler's Method (implicit Euler's Method) (first-order accuracy)

$$X_{n+1} = X_n + hf(t_{n+1}, X_{n+1})$$

$$X_0 = x(t_0) \in R$$

$$n=0, \cdots, N-1$$

(b) Euler's Method: only first-order accurate.
however, it is simple and cheap to implement

Runge-kutta: achive higher accuracy by sacrificing the efficiency of Euler method through re-evaluating $f(\cdot,\cdot)$ at points intermediate between $(t_n, x(t_n))$ and $(t_{n+1}, x(t_{n+1}))$

(second-order) 下面2个方法

① The Modified Euler Method.

$$X_{n+1} = X_n + hf\left(t_n+\frac{h}{2}, X_n+\frac{h}{2}(t_n, X_n)\right)$$

② The Improved Euler Method.

$$X_{n+1} = X_n + \frac{h}{2}\left\{f(t_n, X_n) + f(t_n+h, X_n+hf(t_n, X_n))\right\}.$$

Convergence of implicit method. The implicit method converges much faster that the explicit method since

$$|e_n| \leq \frac{M_3}{12L_\Phi}(e^{L_\phi(t_N-t_0)} - 1)h^2,$$

for some L_ϕ.

Now let us see how we can find L_Φ.To use the error estimation theorem we have to put the implicit method in its standard form, $x_{n+1} = x_n + h\Phi(t_n, x_n; h)$. This yields that for implicit method,

$$h\Phi(t_n, x_n; h) = \frac{h}{2}\left(f\left(t_n, x_n\right) + f\left(t_{n+1}, x_{n+1}\right)\right)$$

$$= \frac{h}{2}(f(t_n, x_n) + f(t_{n+1}, x_n + h\Phi(t_n, x_n; h))$$

Now we use this to find the Lipschitz constant,

$$|\Phi(t_n, u; h) - \Phi(t_n, v; h)|$$

$$= \frac{1}{2}|f(t_n, u) - f(t_n + h, u + h\Phi(t_n, u; h)) - (f(t_n, v) - f(t_n + h, v + hv\Phi(t_n, v; h))|$$

$$\implies |\Phi(t_n, u; h) - \Phi(t_n, v; h)|$$

$$\leq \frac{1}{2}|f(t_n, u) - f(t_n, v)| + \frac{1}{2}|f(t_n + h, u + h\Phi(t_n, u; h) - f(t_n + h, v + h\Phi(t_n, v; h))|$$

$$\leq L_f|u - v| + \frac{1}{2}L_f|u + h\Phi(t_n, u; h) - (v + h\Phi(t_n, v; h))|$$

$$\leq \frac{1}{2}L_f|u - v| + \frac{1}{2}|u - v| + \frac{1}{2}L_f h|\Phi(t_n, u; h) - \Phi(t_n, v; h)|$$

By rearrangement.

$$\left(1 - \frac{1}{2}L_f h\right)|\Phi(t_n, u; h) - \Phi(t_n, v; h)| \leq L_f|u - v|$$

$$\implies |\Phi(t_n, u; h) - \Phi(t_n, v; h)| \leq \frac{L_f}{1 - \frac{1}{2}L_f h}|u - v|$$

$$\implies L_\Phi = \frac{L_f}{1 - \frac{1}{2}L_f h}$$

$$\text{for small } h \Rightarrow L_\Phi = \frac{1}{1-\dfrac{1}{2}} = 2$$

Now using the general estimation,

$$|e_n| \leq \frac{T_{max}}{L_\phi}(e^{L_\Phi(t_N - t_0)} - 1)$$

As $T_{max} = \max_{n=0,..,N-1} T_n \leq \frac{h^2}{12}M_3$. One can see that $e_n \to 0$ as $h \to 0$ at rate h^2.

Recall: geometric Brownian motion

$S_t = S_0 e^{(r-\frac{1}{2}\sigma^2)t+\sigma W_t}$

volatility

Brownian motion.

Stochastic Differential Equation (SDE)

$(S_t)_{t\geq 0}$

4.5. Stochastic differential equation

$dS_t = rS_t dt + \sigma S_t dW_t$.

In this section, we study a dynamic continuous time model which is used to model the dynamic of an stock price and its return. This model is not deterministic for different scenarios, with different probabilities we have different paths. Just reminding that that an ODE, $x' = \frac{dx}{dt} = f(t, x)$, is used to model motion of an object which i not affected by stochastic events. However, in the real world, for example price of stocks, are affected by stochastic events. So if we look at the differential form of an ODE, which is given by $dx = f(t, x)\,dt$, needs an stochastic part. This is not easy, in general. Indeed, to introduce a stochastic increment, we need to introduce a particular type of random process. This random process is called a Brownian motion.

It can be shown that $S=(S_t)_t$ satisfies in the SDE

$\boxed{dS_t = rS_t dt + \sigma S_t dW_t. (*)}$

前面IP都是determined的, 这里是 S(t)

b)

Our aim is to find a discretized numerical approximation for the model. Reminding that, a binomial tree which was studied earlier, is a dynamic model for stock prices. In order to set up the model we need to get familiarized with new concepts.

what is $W=(W_t)_{t\in[0,T]}$?

4.5.1. Brownian motion and stochastic calculus . Denoted by W_t is a continuous time random process where for any scenarios w, the mapping $t \mapsto W_t(w)$ is a continuous function of t and also

(1) $W_0 \equiv 0$

(2) $W_t - W_s \sim W_{t-s} \sim N(0, t-s)$

Gaussian distribution with mean 0 and variance t-s.

equality in distribution

☆ stationary increment.

Independent Increments ☆

⊥ → Independent.

OK

(3) $W_{t_2} - W_{t_1}$ ind. $W_{s_2} - W_{s_1}$, $s_1 < s_2 < t_1 < t_2$.

Remark: Def of Brownian motion: $\begin{cases} E(W_t)=0 \\ Var(W_t - W_s) = t-s \end{cases}$

This process is called a (standard) Brownian motion. We want to model a small changes in the stock prices, that means if we denote the price of the stock with S_t, we want to look at dS_t. As the stock prices have a high level of uncertainty, for each scenario w, $dS_t(w)$ will be different. The best way to do this is to use small changes of the values of a Brownian motion i.e. $dW_t(w)$. But it is not always possible with usual methods, if W_t is not bounded variation. However, one can readily introduce dW_t as $\Delta W_t = W_{t+h} - W_t$ for a very small h. Therefore, if we just accept the legitimacy of the notation, we can model the stock price as,

Itô stochastic Integral $\int_0^T \delta(t, S_t)\,dW_t$.

$$dS_t = \mu(t, S_t)dt + \sigma(t, S_t)dW_t.$$ ☆☆☆

Drift coefficient ↳ volatility coefficient

In this section we briefly review a calculus that is relevant to this type of dynamics. The relevant calculus, that is called *stochastic calculus*, is a rather different calculus. This calculus is very useful in financial math. Let $f(t, x)$ be a function which is smooth enough in both arguments. Let S_t be a random process which follows the following dynamic.

$$dS_t = \mu(t, S_t)dt + \sigma(t, S_t)dW_t.$$

☆ very Important.

The $y_t = f(t, S_t)$ follows a dynamic as follows,

$$dy_t = \left(\frac{\partial f}{\partial t}(t, S_t) + \mu(t, S_t)\frac{\partial f}{\partial x}(t, S_t) + \frac{1}{2}\sigma^2(t, S_t)\frac{\partial^2 f}{\partial x^2}(t, S_t) \right) dt$$
$$+ \sigma(t, S_t)\frac{\partial f}{\partial x}(t, S_t)dW_t$$

To see why this happens we have to know about some general rules.

x	dW_t	dt
dW_t	dt	0
dt	0	0

TABLE 1. Law of stochastic differentials.

Now using this table and the Taylor expansion we have,

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

$(dW_t)(dW_t) = dt$

$(dW_t)(dt) = (dt)(dW_t) = 0 = (dt)(dt)$

Now, consider a smooth function

$(t,x) \mapsto f(t,x)$

look at new stochastic process.

$y_t := f(t, S_t)$, where diffusion S follow dynamic (**).

$f \in C^{1,2}([0,\infty), \mathbb{R})$.

What is the dynamic stochastic process $y = (y_t)$ follows?

Note that, by $\to$ $dS_t = \mu(t,S_t)dt + \sigma(t,S_t)dW_t$.

$$(dS_t)^2 = (\mu(t,S_t)dt + \sigma(t,S_t)dW_t)^2$$

$$= \mu^2(t,S_t)\underbrace{(dt)^2}_{=0} + 2\mu(t,S_t)\sigma(t,S_t)\underbrace{dt \times (dW_t)}_{=0} + \sigma^2(t,S_t)\underbrace{(dW_t)^2}_{=dt}$$

$$= \sigma^2(t,S_t)dt$$

Putting this in previous equation, we get

$$df(t,S_t) = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}(t,S_t)\left(\mu(t,S_t)dt + \sigma(t,S_t)dW_t\right)$$

$$+\frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t,S_t)\sigma^2(t,S_t)dt \quad \text{(Taylor process)}$$

$$= \left(\frac{\partial f}{\partial t}(t,S_t)\right)dt + \frac{\partial f}{\partial x}(t,S_t)\mu(t,S_t) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t,S_t)\sigma^2(t,S_t)$$

$$+\frac{\partial f}{\partial x}(t,S_t)\sigma(t,S_t)dW_t.$$

Example 76. Let X_t be a random process which follows the following dynamic, let $X = (X_t)_{t \geq 0}$ be a stochastic process

$$dX_t = \mu dt + \sigma dW_t$$

where μ, σ are constants. Then $S_t = S_0 e^{X_t}$ follows,

$$dS_t = \left(\mu + \frac{1}{2}\sigma^2\right)S_t dt + \sigma S_t dW_t.$$

Let $f(t,x) = S_0 e^x$, then we need to find the following derivatives:

$$\frac{\partial f}{\partial t} = 0,$$

$$\frac{\partial f}{\partial x} = S_0 e^x,$$

$$\frac{\partial^2 f}{\partial x^2} = S_0 e^x.$$

$\mu \to -\sigma$

Using the Ito lemma we have

$$\frac{\partial f}{\partial t} + \frac{\partial f}{\partial x}\mu + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\sigma^2 = \mu S_0 e^x + \frac{1}{2}\sigma^2 S_0 e^x,$$

and

$$\frac{\partial f}{\partial x}\sigma = \sigma S_0 e^x,$$

and therefore,

$$dS_t = \left(\left(\mu + \frac{1}{2}\sigma^2\right)S_0 e^{X_t}\right)dt + \sigma s_0 e^{X_t} dW_t$$

$$\Rightarrow dS_t = \left(\mu + \frac{1}{2}\sigma^2\right)S_t dt + \sigma S_t dW_t.$$

Example: let $X=(X_t)_{t\geq0}$ be a stochastic process

following dynamic

$$dX_t = \mu dt + \delta dW_t.$$

Def. new stochastic

$$S_t := S_0 e^{X_t} \; ; \; t\geq 0$$

note that, $f(t,x) = S_0 e^x$ ind. of t, Also.

$$\frac{df}{dt} = 0 \; ; \; \frac{df}{dx} = S_0 e^x \; ; \; \frac{d^2 f}{dx^2} = S_0 e^x.$$

$\mu(t)$ $\quad$ $\mu(t,S_t)$

$$df(t,S_t) = \left(\frac{\partial f}{\partial t}(t,S_t)\right) dt + \frac{\partial f}{\partial x}(t,S_t)\mu(t,S_t) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t,S_t)\sigma^2(t,S_t)$$
$$+ \frac{\partial f}{\partial x}(t,S_t)\sigma(t,S_t)dW_t.$$

(plug in)

$\Rightarrow$ Itô's Lemma / formula. formula

$$dS_t = df_t = \left(S_0 \mu e^{X_t} + \frac{1}{2} S_0 \delta^2 e^{X_t}\right)dt + S_0 \delta e^{X_t} dW_t.$$
$$= \left(\mu + \frac{1}{2}\delta^2\right) S_t \, dt + \delta S_t dW_t.$$

Remark:

Let $dS_t = \mu(t,S_t)dt + \delta(t,S_t)dW_t$

Assume $f \in C^{1,2}([0,\infty)\times\mathbb{R})$

$(t,x) \mapsto f(t,x)$

and let $y_t := f(t,S_t)$. Then

tell us that $y=(y_t)_{t\geq0}$ satisfies in dynamic

$$dy_t = \left(\frac{\partial f}{\partial t}(t,S_t)\right) dt + \frac{\partial f}{\partial x}(t,S_t)\mu(t,S_t) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t,S_t)\sigma^2(t,S_t)$$
$$+ \frac{\partial f}{\partial x}(t,S_t)\sigma(t,S_t)dW_t.$$

and therefore,

$$dS_t = \left(\left(\mu + \frac{1}{2}\sigma^2 \right) S_0 e^{X_t} \right) dt + \sigma s_0 e^{X_t} dW_t$$

$$\Rightarrow dS_t = \left(\mu + \frac{1}{2}\sigma^2 \right) S_t dt + \sigma S_t dW_t.$$

EXAMPLE 77. If r_t follows,

$$dr_t = r dt + \sigma dW_t$$

What is the dynamic of $R_t = r_t^2$? Let $f(t,x) = x^2$, we have to the following derivatives

$$\frac{\partial f}{\partial t} = 0, \frac{\partial f}{\partial x} = 2x \text{ and } \frac{\partial^2 f}{\partial x^2} = 2.$$

So we have

$$\frac{\partial f}{\partial t} + \mu \frac{\partial f}{\partial x} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\sigma = 2rx + \sigma^2$$

$$\frac{\partial f}{\partial x} = 2x \Longrightarrow \frac{\partial f}{\partial x}\sigma = 2x\sigma,$$

and using the Ito lemma we have:

$$dR_t = (2rr_t + \sigma^2)dt + 2\sigma r_t dW_t.$$

EXAMPLE 78. Show $S_t = (W_t + \sqrt{s_0})^2$ solves,

$$S_t = t + s_0 + 2\int_0^t \sqrt{s} dW_s$$

This is equivalent to $dS_t = dt + 2\sqrt{S_t} dW_t$. Let $X_t = W_t \Longrightarrow \mu = 0$, $\sigma = 1$ and $f(t,x) = (x + \sqrt{s_0})^2$, then,

$$\frac{\partial f}{\partial t} = 0,$$

$$\frac{\partial f}{\partial x} = 2(x + \sqrt{s_0}),$$

$$\frac{\partial^2 f}{\partial x^2} = 2.$$

Observe,

$$\frac{\partial f}{\partial t} + \mu\frac{\partial f}{\partial x} + \frac{1}{2}\sigma^2\frac{\partial^2 f}{\partial x^2} = 1$$
$$\frac{\partial f}{\partial x}\sigma = 2(x - \sqrt{s_0}).$$

Using the Ito lemma we get

$$dS_t = dt + 2(W_t + \sqrt{s_0})dW_t$$

but $W_t + \sqrt{s_0} = \sqrt{s_t} = \sqrt{(W_t + \sqrt{s_0})^2}$.

4.6. SDE Discretization

Yamada-Watanabe theory.

Let us focus on discretization of a SDE. Let us split the time interval $[t_0, T]$ to N sub-intervals:

consider a diffusion.

$[t_n, t_{n+1}]$; $n = 0, 1, \dots, N-1$

$$t_n = t_0 + nh, n = 0, ..., N$$

↳ time discretization points

(*) $dS_t = \mu(t, S_t)dt + \delta(t, S_t)dW_t$.

⇒ stochastic dynamic

where

$$h = \frac{T - t_0}{N}.$$

(step size) mesh size ←

$S_t = S_0 + \int_0^t \mu(s, S_s)dt + \int_0^t \delta(s, S_s)dW_s$

Initial value. ordinary intergral stochastic Itô integral)

We approximate the differentiation of W_t, dW_t by, a small increment as follows

$$dW_t \approx \Delta W_t := W_{t+h} - W_t$$

As increments are i.i.d then,

how Ito's apply in stchastic dynamic of $f(t, S_t)$

$$\Delta W_t \sim N(0, h)$$

$S = (S_t)_{t \geq 0}$ ⇒ Ito's process.

and ΔW_t are independent. Therefore, if S_t follows,

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

$(t, x) \mapsto f(t, x) \in C^{1,2}(\mathbb{R}_t, \mathbb{R})$

denote the space of all function in 2 variables $(t, x$.

then the discrete version of the SDE is given by

$$dS_t \approx \Delta S_t = \mu(t, S_t)h + \sigma(t, S_t)\Delta W_t.$$

If $h \to 0$ then ΔS_t converges to real dS_t in a "particular scene". As for the integral we can sum up the differentiation, i.e,

$$\int_0^T f(t, S_t)dS_t \approx \sum_{t=0,h,...,T} f(t, S_t)\Delta S_t.$$

Note that $f(t, S_t)$ ind. ΔS_t.

当 $\delta = 0$, 它 reduce to O.

The Euler-Maruyama discretization:

$X = (X_t)_{t \in [t_0, T)}$.

satisfying the iterative scheme.

(**) $X_{n+1} = X_n + \mu(t_n, X_n)h + \delta(t_n, X_n)\Delta W_{t_n}$

(**) $X_{n+1} = X_n + a(t_n, X_n)h + b(t_n, X_n)\Delta W_n$

where $h = t_{n+1} - t_n$ is the step size

$\Delta W_{t_n} := W_{t_{n+1}} - W_{t_n} = W_{t_{n+1}} - W_{t_n}$, for $n = 0, 1, \dots N-1$, with initial value $Y_0 = S_0$.

where we have writen $X_n = X_{t_n}$.

Important note: In order to simulate different paths we can consider an i.i.d., sequence $\{\epsilon_t\}_{t=1}^{N}$ of $N(0,1)$, and put $\Delta W_t = \sqrt{h}\epsilon_t$.

EXAMPLE 79. Consider an SDE $dS_t = \mu(t, S_t)dt + \sigma(t, S_t)dW_t$, for three

1) A Brownian motion with $\mu(t,x) = 0, \sigma(t,x) = 1$;
2) A geometric Brownian motion with drift 0.05 and volatility 0.1, with $\mu(t,x) = 0.05x, \sigma(t,x) = 0.1x$;
3) An Orstein-Uhlenbeck process with $\mu(t,x) = (1-x), \sigma(t,x) = 0.25$.

By discretization, the processes become

$$S_{(i+1)h} = S_{ih} + \Delta W_t;$$
$$S_{(i+1)h} = S_{ih} + 0.05 S_{ih} h + 0.1 S_{ih} \Delta W_t;$$
$$S_{(i+1)h} = (1 - S_{ih})\, h + 0.25 \Delta W_t.$$

Now consider $T = 1$, $h = 0.001$. In the following we show the simulated paths for 10 paths.

Note also that the sequence

$\{X_n : n = 0, 1, \dots N-1\}$.

of value of the Euler-Maruyama scheme (**)

可以以一个相似的方式计算.

The main difference:

$\boxed{\Delta W_n := W_{t_{n+1}} - W_{t_n}}$ (***)

for $n = 0, 1, \dots, N-1$ of the Brownian motion

$W = (W_t)_{t \in [t_0, T]}$

But, remember that due to properties of Brownian motion, the increment (***) are i.i.d. with

$\Delta W_n \sim N(0, h)$

and so $\forall n = 0, 1, \dots N-1$

$E(\Delta W_n) = 0$

$E((\Delta W_n)^2) = h$.

$t_n \le t \le t_{n+1}$. $\quad X_t = X_{t_n} + \dfrac{t - t_n}{t_{n+1} - t_n}(X_{t_{n+1}} - X_{t_n})$

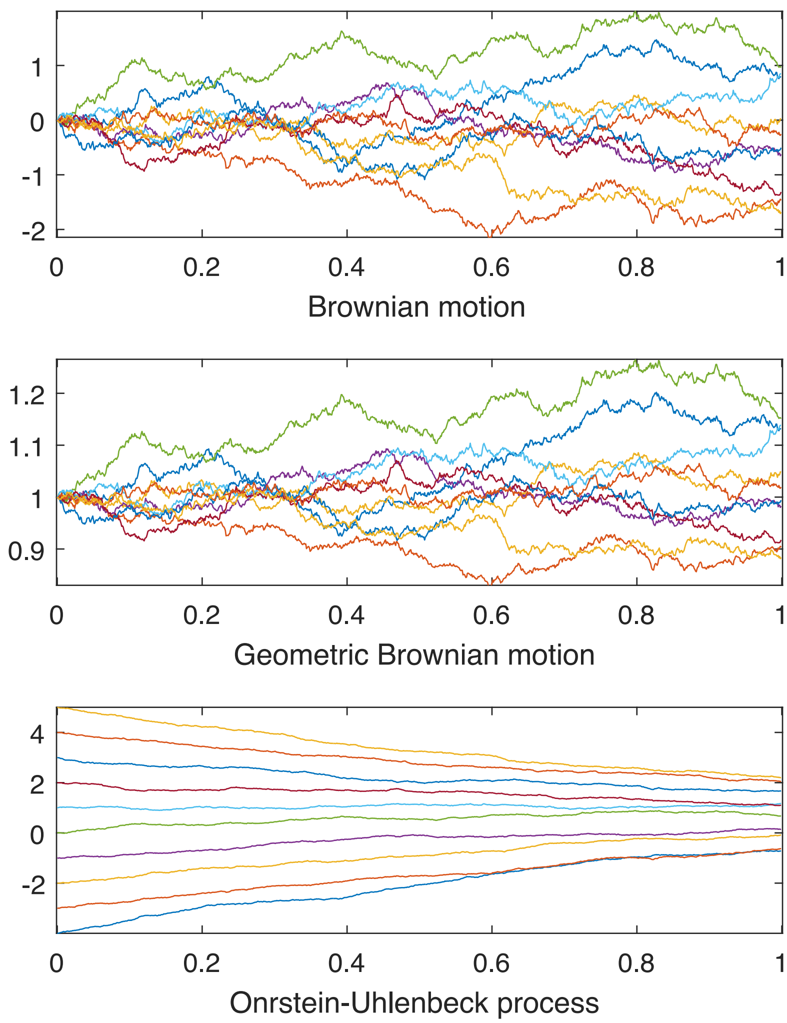


FIGURE 4.6.1. Ten sample paths.

EXAMPLE 80. Let us see what is the discretization of,

→ recall !!!

$dS_t = \mu(t, S_t)dt + \sigma(t, S_t)dW_t.$

$S_t = S_0 + \int_0^t m(s, S_s)dt + \int_0^t \sigma(s, S_s)dW_s$

$(**)\ X_{n+1} = X_n + m(t_n, X_n)h + \sigma(t_n, X_n)\Delta W_{t_n}.$

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

The discretization of this for is easily given by,

$$\Delta S_t = S_{t+h} - S_t = \mu h + \sigma \left(W_{t+h} - W_t\right)$$

Euler discretization used that.

EXAMPLE 81. With discretization we want to find the variance of the following process, S_t,

$$\begin{cases} dS_t = \sigma\left(t, S_t\right) dW_t, \\ s_0 = 0. \quad (t_0 = 0) \end{cases}$$

solution is given by Itô stochastic integral. $S_t = \int_0^t \sigma(u, S_u)dW_u;\ t \in [0, T]$

$$\Delta S_t = S_{t+h} - S_t = \sigma(t, S_t)(W_{t+h} - W_t) = \sigma(t, S_t)\Delta W_t.$$

where $\sigma(t, S_t)\Delta W_t$ are independent. Therefore, we have $S_T = \int_0^T \sigma(t, S_t)dW_t$ and

$$E\left(S_T\right) = E\left(\int_0^T \sigma(t, S_t)dW_t\right).$$

The Euler-Maruyama approximation is given by $X_{n+1} = X_n + \sigma(t_n, X_n)\Delta W_n;\ n = 0, 1, \dots N-1$, where as before $X_n = X_{t_n}$. $t_n = t_0 + nh = nh$. $h = \frac{T}{n}$.

$S_t = \sum_{k=0}^{m-1} \sigma(t_k, S_{t_k})\Delta W_k.$

Observe,

$$S_t \approx \sum_{i=0}^{N-1} \sigma\left(t_i, S_{t_i}\right) \Delta W_{t_i}$$

$$\Rightarrow E(S_t) \approx \sum_{i=0}^{N-1} E\left(\sigma(t_i, S_{t_i})\Delta W_{t_i}\right)$$

$$= \sum_{i=0}^{N-1} E\left(\sigma\left(t_i, S_{t_i}\right)\right) \underbrace{E\left(\Delta W_{t_i}\right)}_{=0} = 0.$$

independent

$\Delta W_t \sim N(0, h)$

Therefore, $E(S_t) = 0$ for every discretization point t, and by using an approximation argument as step size $h \to 0$ (or equivalently $N \to \infty$) we can deduce that "= 重偏"

On the other hand, we have,

$$S_t^2 = \left(\sum_{i=0}^{N-1} \sigma(t_i, S_{t_i})\Delta W_{t_i}\right)^2$$

are independent

$$= \sum_{i=0}^{N-1} \sigma\left(t_i, S_{t_i}\right)^2 \left(\Delta W_{t_i}\right)^2$$

$$+2 \sum_{i<j} \sigma\left(t_i, S_{t_i}\right) \sigma\left(t_j, S_{t_j}\right) \Delta W_{t_i} \Delta W_{t_j},$$

where $\sigma(t_i, S_{t_j})^2$ and $(\Delta W_{t_i})^2$ are independent. As a result,

$$E\left(S_t^2\right) = \sum_{i=0}^{N-1} E\left(\sigma^2\left(t_i, S_{t_i}\right)\left(\Delta W_{t_i}\right)^2\right)$$

$$+2\sum_{i<j} E\left(\sigma\left(t_i, S_{t_i}\right)\sigma\left(t_j, S_{t_j}\right)\Delta W_{t_i}\Delta W_{t_j}\right)$$

$$= \sum_{i=0}^{N-1} E\left(\sigma^2\left(t_i, S_{t_i}\right)\right) E\left(\Delta W_{t_i}\right)^2 \rightarrow \frac{T}{N}.\ \frac{T}{n}.$$

Recall:

$$\begin{cases} E(\Delta W_k)=0 \\ E((\Delta W_k)^2)=h=\frac{T}{n} \end{cases}$$

$$+2\sum_{i<j} E\left(\sigma\left(t_i, S_{t_i}\right)\sigma\left(t_j, S_{t_j}\right)\right)\underbrace{\Delta W_{t_i}\Delta W_{t_j}}_{=0}$$

$$= \sum_{i=0}^{N-1} \overset{\frac{T}{N}.}{h} E\left(\sigma^2\left(t_i, S_{t_i}\right)\right) = \frac{1}{N}\sum_{i=0}^{N-1} E\left(\sigma^2\left(t_i, S_{t_i}\right)\right).$$

If $N \to \infty$, then, $\int_0^T E(\sigma^2(t, S_t))dt$. Therefore,

$$\frac{T}{N}\sum_{i=0}^{N-1}\cdots = \int_0^T \cdots dt$$

$$Var(S_T) = \int_0^T E\left(\sigma^2\left(t, S_t\right)\right) dt.$$

$$\Rightarrow E(S_t^2) = Var(S_t) = \int_0^t E(\sigma^2(u, S_u))du \quad \forall t \in [0, T].$$

In the particular case $\sigma(t,x) = \sigma(t)$ is just a function of time, we can infer that $S_t = \int_0^t \sigma(u)\,dW_t \sim N(0, \int_0^t \sigma^2(u)\,du)$

Gaussian distributed with zero mean and variance $\int_0^t \sigma^2(u)du$.

$W_t \sim N(0, t)$

$\Rightarrow X_t = W_t$

$dX_t = dW_t \Rightarrow \mu=0, \sigma=1$

Example: Find the dynamic of $\frac{W_t^2}{2}$, where $W=(W_t)_{t\in[0,T]}$ is a $(0,1)$ Brownian motion.

We use Itô formula with function

$$f(t,x)=\frac{x^2}{2}.$$

① note that $\frac{\partial f}{\partial x}=x$, and $\frac{\partial^2 f}{\partial^2 x}=1$.

$$f(t,S_t)=\left(\frac{\partial f}{\partial t}(t,S_t)\right)dt+\frac{\partial f}{\partial x}(t,S_t)\mu(t,S_t)+\frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t,S_t)\sigma^2(t,S_t)$$
$$+\frac{\partial f}{\partial x}(t,S_t)\sigma(t,S_t)dW_t.$$

$$=\left(\mu x+1\times\frac{1}{2}\sigma^2\right)dt+\sigma x\,dW_t$$

hence

$$d(f(t,W_t))=d\left(\frac{W_t^2}{2}\right)=W_t dW_t+\frac{1}{2}dt$$

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

whereas in the deterministic case

$$\int_0^t X_s dX_s=\frac{X_t^2}{2} \quad X_0=0$$

$$E\left(\int_0^t W_s dW_s\right)=E\left(\frac{W_t}{2}-\frac{t}{2}\right)=\frac{1}{2}E(W_t^2)-\frac{t}{2}=\frac{t}{2}-\frac{t}{2}=0$$

which is consistance with the previous example.

and $\mathrm{Var}\left(\int_0^t W_s dW_s\right)^2=\int_0^t E(W_s^2)ds=\int_0^t s\,ds=\frac{t^2}{2}$

that is the same as

$W_t\sim N(0,t)$

$\Uparrow \quad E(W_t^4)=3t^2$

$$\triangle\ \mathrm{Var}\left(\frac{W_t^2}{2}-\frac{t}{2}\right)=\mathrm{Var}\left(\frac{W_t^2}{2}\right)=E\left(\left(\frac{W_t^2}{2}\right)^2\right)-\left(\frac{t}{2}\right)^2=\frac{1}{4}E(W_t^4)-\frac{t^2}{4}=\frac{3t^2}{4}-\frac{t^2}{4}=\frac{t^2}{2}$$

$$\int_{-\infty}^{\infty}\frac{x^4}{\sqrt{2\pi t}}e^{-\frac{x^2}{2t}}dx=\frac{1}{\sqrt{2\pi t}}\int_{-\infty}^{\infty}x^4e^{-\frac{x^2}{2t}}dx=3t^2$$ （没算出来）

Example:

(a) Riemann Integral:

Let $f: [a,b] \to \mathbb{R}$ be a bounded function.

Then we say that f is Riemann-Integrable.

if the following limit exists:

$$(*)\ \int_a^b f(t)\,dt = \lim_{\|\Pi_n\|\to 0} \sum_{k=1}^{n} f(\gamma_i)(t_i - t_{i-1})$$

where have $\Pi_n = \{a = t_0 < t_1 \cdots t_n = b\}$ is a partition of $[a,b]$

$\|\Pi_n\| = \max_{1\le i\le n} (t_i - t_{i-1})$ is the partition. mesh.

and $\gamma_i \in [t_{i-1}, t_i]$ is an evaluation point.

(b) Riemann-Stieltjes Integral:

Let g be an increasing function on $[a,b]$.

A bounded function $f: [a,b] \to \mathbb{R}$ is said to be

Riemann-Stieltjes Integral:

g:

$$\int_a^b f(t)\,dg(t) = \lim_{\|\Pi_n\|\to 0} \sum_{k=1}^{n} f(\tau_i)(g(t_i) - g(t_{i-1}))$$

$\tau_i [t_{i-1}, t_i]$ can be any evaluation point.

Again, we know that if f is a continuous function

on $[a,b]$.

(c) Stochastic Integral.

Let $W = (W_t)_{t\in[0,T]}$ be a Brownian motion

Let $L_n: \sum_{k=1}^{n} W_{t_{i-1}} (W_{t_i} - W_{t_{i-1}})$ ← $\tau_i = t_{i-1}$ is the left point of interval $[t_{i-1}, t_i]$.

$R_n: \sum_{k=1}^{n} W_{t_i} (W_{t_i} - W_{t_{i-1}})$.

$\tau_i = t_i$.

Then $\begin{cases} L_n \xrightarrow{\|\Pi_n\|\to 0} \frac{1}{2}W_T^2 - \frac{T}{2} \\ R_n \xrightarrow{\|\Pi_n\|\to 0} \frac{1}{2}W_T^2 + \frac{T}{2}, \end{cases}$ where the limits are taken in $L^2(\Omega)$ sense, and $\Pi_n = \{0, t_0 < t_1 < \cdots t_n = T\}$ is a partition of $[0, T]$

$\tau_i \in [t_{i-1}, t_i]$.

particular choice $\rightarrow$ $\tau_i = \frac{t_i + t_{i-1}}{2}$.

(mid point

① Itô stochastic Integral ($\tau_i = t_{i-1}$)

(Itô) $\int_0^T W_t \, dW_t = \frac{1}{2} W_T^2 - \frac{1}{2} T$

② stratonovich stochastic Integral ($\tau_i = \frac{t_i + t_{i-1}}{2}$).

(Stratonovich) $\int_0^T W_t \, dW_t = \frac{1}{2} W_T^2$.

$\int_0^T W_t \circ dW_t$.

(unlike Itô calculus) it satisfies in the chain rule

$\int_0^T f'(W_t) \circ dW_t = f(W_T) - f(W_0)$

4.7. Exercise

EXERCISE 82. Verify that the following functions satisfy a Lipschitz condition on the respective intervals and find the associated Lipschitz constants:
a) $f(t,x) = 2xt^{-4}, \quad t \in [1,\infty);$
b) $f(t,x) = 2x(1+x^2)^{-1}(1+e^{-|t|}), \qquad t \in (-\infty,\infty);$

EXERCISE 83. Consider the following ODE

$$\begin{cases} x^2 y + \frac{1}{2} y' = 0 & 0 \le x \le T \\ y(0) = 1 \end{cases},$$

for some positive number T.

a) Write the problem in its general form.

b) Show that this problem has a global solution.

c) Write the explicit and implicit numerical schemes for this ODE.

d) If possible find an explicit answer to this ODE.

EXERCISE 84. Let us assume y is a function of x and consider the same ODE in the last question

$$\begin{cases} x^2 y + \frac{1}{2} y' = 0 & x \ge 0 \\ y(0) = 1 \end{cases}.$$

Let $f(x) = \int_0^x y(s)\,ds$. Now introduce $S_t = f(X_t)$ where X_t dynamic is given as follows

$$dX_t = X_t^2 dt + dW_t.$$

a) Find the SDE for S_t then use it to give a discrete form of the SDE.

b) Use the SDE and find discretization of S_t

c) Find $E(S_T)$ and $Var\,(S_T)$.

EXERCISE 85. Let us assume that the return of a stock is denoted by r and its dynamic is given as

$$dr_t = \sin(t)dt + \sigma dW_t, r_0 = 0,$$

where $\sigma > 0$ is a real number and W_t is a Brownian motion.

a) Using Itô formula, find a dynamic for $\exp(r_T)$

b) Find the discretization of r_T.

c) Use the discretization in b) to find $E(\exp(r_T))$ and send $N \to \infty$.

d) Find $E(\exp(r_T))$ from part a) and show it is equal to what you have find in limit in c).

EXERCISE 86. Let us assume that the return of a stock is denoted by r and its dynamic is given as

$$\begin{cases} dX_t = \mu\,(t, X_t)\,ds + \sigma\,(t, X_t)\,dW_t, 0 \leq t \leq T, \\ X_0 \in R \end{cases}$$

where $\mu\,(t, x)$ and $\sigma\,(t, x) > 0$ are real functions and $(W_s)_{s\geq 0}$ is a Brownian motion. Let us fix a natural number $N \in \mathbb{N}$ and define $t_i = (i/N)t$, $i = 0, ..., N$. Answer the following questions

a) Find the discretization of X_t and use it to find an ODE for $x_t = E(X_t)$ and show

$$\begin{cases} x' = E\,(\mu\,(t, X_t)) \\ x_0 = X_0 \end{cases} .$$

Example: 85

EXERCISE 85. Let us assume that the return of a stock is denoted by r and its dynamic is given as

$$dr_t = \sin(t)dt + \sigma dW_t, r_0 = 0,$$

where $\sigma > 0$ is a real number and W_t is a Brownian motion.

a) Using Itô formula, find a dynamic for $\exp(r_T)$

b) Find the discretization of r_T. Find Euler-Maruyama discretization.

c) Use the discretization in b) to find $E(\exp(r_T))$ and send $N \to \infty$.

d) Find $E(\exp(r_T))$ from part a) and show it is equal to what you have find in limit in c).

(a) $f(t,x) = \exp(x)$

$$\frac{\partial f}{\partial t} = 0, \quad \frac{\partial f}{\partial x} = \exp(x), \quad \frac{\partial^2 f}{\partial x^2} = \exp(x)$$

Thus, Itô's formula implies that

$$Y_t := f(t, r_t) = \exp(r_t)$$

$$dY_t = d(\exp(r_t))$$

$$= (\sin(t)\exp(r_t) + \tfrac{1}{2}\sigma^2 \exp(r_t))dt + \sigma \exp(r_t) dW_t$$

$$= Y_t(\sin(t) + \tfrac{1}{2}\sigma^2)dt + \sigma Y_t dW_t.$$

(b) Let $0 = t_0 < t_1 < \cdots < t_N = T$, when $t_n = nh$.

with $h = \frac{T}{N}$ be a uniform partition of $[0,T]$.

Then $X_{n+1} = X_n + \sin(t_n)h + \sigma \Delta W_n$ → stochastic (random) noise input.

$X_0 = 0$

$n = 0, 1, \cdots N-1$ and as before $\Delta W_n = W_{t_{n+1}} - W_{t_n}$.

what is important is that the random noise inputs $\Delta W_n \sim N(0,h)$ are i.i.d. with

Gaussian distribution (zero mean and variance h).

(c). Use discretization to find $E(\exp(r_T))$ by sending $N \to \infty$.

Note that first we can write.

$$r_T = \sum_{k=0}^{N-1} \Delta r_{\frac{kT}{N}} \quad (\text{Telescopic sum}).$$

$$= \sum_{k=0}^{N-1} \left(\sin\left(\tfrac{kT}{N}\right) h + \delta \Delta W_{\frac{kT}{N}} \right).$$

Hence,

$$E(\exp(r_T)) = E\left(\exp\left(\sum_{k=0}^{N-1} \left(\sin\left(\tfrac{kT}{N}\right) h + \delta \Delta W_{\frac{kT}{N}}\right)\right)\right)$$

$$= \exp\Big(\underbrace{\sum_{k=0}^{N-1} \sin\left(\tfrac{kT}{N}\right) h}_{\text{determined,}}\Big) E\Big(\exp\Big(\underbrace{\sum_{k=0}^{N-1} \delta \Delta W_{\frac{kT}{N}}}_{\text{Telescopic sum.}}\Big)\Big).$$

拿出 E

$$= \exp\left(\sum_{k=0}^{N-1} \sin\left(\tfrac{kT}{N}\right) h\right) E(\delta W_T)$$

Riemann sums converge to Riemann integral

as $N \to \infty$ $\quad \exp\left(\int_0^T \sin(t)dt\right) \exp\left(\frac{\delta^2 T}{2}\right).$

$(h \to 0)$

or equivallanty

$$= \exp\left(1 - \cos(T) + \frac{\delta^2 T}{2}\right)$$

(d) d) Find $E(\exp(r_T))$ from part a) and show it is equal to what you have find in limit in c).

In part (a). $y_t = \exp(r_t)$

$$dy_t = y_t\left(\sin(t) + \tfrac{1}{2}\delta^2\right) dt + \delta y_t dW_t.$$

In other words (note that $y_0 = 1$)

$$y_t = 1 + \int_0^t y_u \left(\sin(u) + \tfrac{1}{2}\delta^2\right) du + \int_0^t \delta y_u dW_u.$$

Now, take expectation from both sides, and note that the Itô stochastic integrals zero expectation; $E\left(\int_0^t \sigma y_u \, dW_u\right) = 0$

$E[S_t] = E\left[\int_0^t \sigma(u, S_u) \, dW_t\right] = 0 \quad \forall t \in [0,T]$

The latter property means that Itô stochastic integral has zero expectation.

We obtain that

$$E[y_t] = 1 + E\left(\int_0^t y_u (\sin(u) + \tfrac{1}{2}\sigma^2) \, du\right)$$
$$= 1 + \int_0^t E[y_u] (\sin(u) + \tfrac{1}{2}\sigma^2) \, du$$

Denote $Z_t = E[y_t]$. The last equation can be written as

$$\begin{cases} Z_t = 1 + \int_0^t Z_u (\sin(u) + \tfrac{1}{2}\sigma^2) \, du \\ Z_0 = 1 \end{cases} \Rightarrow Z_t' = \left(1 + \int_0^t Z_u (\sin(u) + \tfrac{1}{2}\sigma^2) \, du\right)'$$

i.e. $Z_t' = Z_t (\sin(t) + \frac{1}{2}\sigma^2)$

(from part (c))

Hence. $\dfrac{Z_t'}{Z_t} = \sin(t) + \frac{1}{2}\sigma^2$

And Notice $Z_t = E[y_t] = E[\exp(Y_t)] = \exp\left(1 - \cos(t) + \frac{\sigma^2 t}{2}\right)$

So $Z_t = \exp\left(1 - \cos(t) + \frac{\sigma^2 t}{2}\right)$

$\log(Z_t) = 1 - \cos(t) + \frac{\sigma^2 t}{2}$

On the other hand

$$\frac{Z_t'}{Z_t} = \frac{d(\log Z_t)}{dt}$$

and $\dfrac{d(\log(Z_t))}{dt} = \sin(t) + \frac{1}{2}\sigma^2$.

So.

Therefore, taking integral over $[0,T]$, we obtain $\log(Z_T) - \underbrace{\log(Z_0)}_{=0} = \int_0^T (\sin(t) + \frac{1}{2}\sigma^2) \, dt$

$= 1 - \cos(T) + \frac{1}{2}\sigma^2 T$

$\Rightarrow Z_T = E[y_T] = E[\exp(Y_T)] = \exp\left(1 - \cos(T) + \frac{1}{2}\sigma_T^2\right)$

Now, take expectation from both sides, and note that the Itô stochastic integrals zero expectation; $E\left(\int_0^t \sigma y_u \, dW_u\right) = 0$

$E[S_t] = E\left[\int_0^t \sigma(u, S_u) dW_t\right] = 0 \quad \forall\, t \in [0, T]$

The latter property means that Itô stochastic integral has zero expectation.

We obtain that

$$E[y_t] = 1 + E\left(\int_0^t y_u (\sin(u) + \tfrac{1}{2}\sigma^2) \, du\right)$$
$$= 1 + \int_0^t E[y_u] (\sin(u) + \tfrac{1}{2}\sigma^2) \, du$$

Denote $Z_t = E[y_t]$. The last equation can be written as

$$\begin{cases} Z_t = 1 + \int_0^t Z_u (\sin(u) + \tfrac{1}{2}\sigma^2) \, du \\ Z_0 = 1 \end{cases} \Rightarrow Z_t' = \left(1 + \int_0^t Z_u (\sin(u) + \tfrac{1}{2}\sigma^2) \, du\right)'$$

i.e. $Z_t' = Z_t (\sin(t) + \tfrac{1}{2}\sigma^2)$

(from part (c))

Hence, $\dfrac{Z_t'}{Z_t} = \sin(t) + \tfrac{1}{2}\sigma^2$

And Notice $Z_t = E[y_t] = E[\exp(r_t)] = \exp\left(1 - \cos(t) + \frac{\sigma^2 t}{2}\right)$

So $Z_t = \exp\left(1 - \cos(t) + \frac{\sigma^2 t}{2}\right)$

$\log(Z_t) = 1 - \cos(t) + \frac{\sigma^2 t}{2}$

On the other hand

$$\frac{Z_t'}{Z_t} = \frac{d(\log Z_t)}{dt}$$

and $\dfrac{d(\log(Z_t))}{dt} = \sin(t) + \tfrac{1}{2}\sigma^2$.

So.

Therefore, taking integral over $[0, T]$, we obtain $\log(Z_T) - \underbrace{\log(Z_0)}_{=0} = \int_0^T (\sin(t) + \tfrac{1}{2}\sigma^2) \, dt$

$= 1 - \cos(T) + \tfrac{1}{2}\sigma^2 T$

$\Rightarrow Z_T = E[y_T] = E[\exp(r_T)] = \exp\left(1 - \cos(T) + \tfrac{1}{2}\sigma_T^2\right)$

b) Use the Itô formula to find the SDE of $Z_t = X_t^2$. Find its discretization and use it to find $E(X_t^2)$ and show

$$\begin{cases} z' = E\left(\left(2X_t\mu\left(t, X_t\right) + \sigma^2\left(t, X_t\right)\right)\right) \\ z_0 = X_0^2 \end{cases}.$$

c) Now consider the following cases:

(1) For, $dX_t = -X_t dt + dW_t$, show

$$\begin{cases} x' = -x \\ x_0 = X_0 \end{cases},$$

and

$$\begin{cases} z' = -2z + 1 \\ z_0 = X_0^2 \end{cases}.$$

(2) For $dX_t = (1 - X_t)ds + dW_t$, show

$$\begin{cases} x' = 1 - x \\ x_0 = X_0 \end{cases},$$

and

$$\begin{cases} z' + 2z = 2e^{-t}(X_0 - 1) + 3 \\ x_0 = X_0^2 \end{cases}.$$

(3) For $dX_t = (1 - X_t)\,dt + \sqrt{X_t}dW_t$, show

$$\begin{cases} x' = 1 - x \\ x_0 = X_0 \end{cases},$$

and

$$\begin{cases} z' + 2z = 3\left(1 + e^{-t}\left(X_0 - 1\right)\right) \\ z_0 = X_0^2 \end{cases}.$$

Exercise 87. Let us consider the following SDE

$$dX_t = (1 - X_t)dt + \sqrt{X_t}dW_t, X_0 > 0,$$

where W_t is a standard Brownian motion.

1) If we assume always $x_t > 0$, use the Itô calculus to find a SDE that $S_t = \frac{1}{x_t}$ satisfies.

2) By discretization find an ODE that $\bar{z}_t = E(S_t)$ satisfies.

CHAPTER 5

Computer programs

5.1. Chapter 1 programs

5.1.1. Bisection. This program uses the bisection method.

```
clc
clear
f=@(x)(1/(1+x)-exp(x)+1); % Function f

a=0; % Initial left point
b=.6; % Initial right point
delta=.0000001; % Accuracy

A=[0 0 0]; % Left-mid-right values of the function
while(abs(f(a))>delta) % Condition to reach the accuracy

    c=(a+b)/2; % Finding mid point
    A=[A; [f(a) f(c) f(b)]]; % Left-mid-right values of the function

    if f(a)*f(c)<0 %Check if the solution is in the left interval
        b=c;
    else
        a=c;
    end

end

format long
xlswrite('Bisec.xlsx',A) % Write results in an EXCEL file
A % Display the values of the function
c % Display the approximated root
```

5.1.2. Fixed point.

This program uses a fixed point method.

```
clc
clear
f=@(x)(1/(1+x)-exp(x)+1); % Function f

% the derivative of the function f is ff=@(x)(-1./(1+x).^2-exp(x)). This is
% helpful to find a lambda close to the 1/ff(r). Here we use lambda=2,
% which is close enough to this value.

g=@(x)(x-(1/(-2))*f(x));
%The fixed point of function g is equal to the root of f


delta=.0000001; %Accuracy 10^-7
x0=0.6; %Starting point
r=x0;

while(abs(g(x0)-x0)>delta) % This is the condition of the loop
    x0=g(x0); % Updating
    r=[r;x0]; % Recording all values found in the iteration
end

xlswrite('Fixed.xlsx',r) % Writing in EXCEL

r % Displaying the results
```

5.1.3. Newton method. This program uses a Newton method .

```
clc
clear
f=@(x)(1/(1+x)-exp(x)+1); % Function f
ff=@(x)(-1./(1+x).^2-exp(x)); % Derivative of f

delta=.0000001; % Accurcy
x0=0.6; % Initial point
x=x0;

while(abs(f(x0))>delta)

    x0=x0-f(x0)/ff(x0); % Updating

    x=[x;x0]; % Recording all values found in the iteration

end

xlswrite('Newton.xlsx',x) % Writing in an EXCEL file

x %Displaying the results
```

5.1.4. Secant method. This program uses a Secant method.

```
clc
clear
f=@(x)(1/(1+x)-exp(x)+1); % Function f

delta=.0000001;
n=1;
x0=0.6; % First initial value
x1=0.7; %Second initial value
x=x0;

while(abs(f(x0))>delta)

    xx=x1-(x1-x0)*(f(x1)/(f(x1)-f(x0))); %Updating
    x0=x1;
    x1=xx; % Updating

    x=[x;x0];

end

xlswrite('Secant.xlsx',x) % Writing in an Excel file

x % Displaying approximated roots
```

5.1.5. Jacobi method. This program uses a Jacobi method.

```
clear
clc
n=3; % The dimention of the matrix
x=ones(n,1);
b=x;  % Vector b
y=x';

for i=1:n
    for j=1:n
        A(i,j)=exp(-2*abs(i-j)); % Matirx of coeficients
    end
end

LU=A-eye(n,n); % LU=L+U
g=@(x)(-LU*x+b); % The fixed point function

while(max(abs(g(x)-x))>0.0000001) % Condition to reach accuracy

    x=g(x); % Updating
    y=[y;x']; % Recording the values in the iterations
end
xlswrite('Jacobi.xlsx',y); % Writing in an EXCEL file
y %Display the values in the iterations
```

5.1.6. Gauss-Sieldel method.

This program uses a Gauss-Siedel method.

```
clear
clc

b=x; % Vector b
n=3; % The dimention of the matrix
x=ones(n,1);

for i=1:n
    for j=1:n
        LD(i,j)=(i<=j)*exp(-2*abs(i-j));  % LD=L+D
        A(i,j)=exp(-2*abs(i-j));% Matirxx of coeficients
    end
end

U=A-LD;
y=x';
g=@(x)(-inv(LD)*U*x+b); % The fixed point funciton

while(max(abs(g(x)-x))>0.0000001)
    x=g(x);
    y=[y;x'];
end
xlswrite('Gauss.xlsx',y)
y % Display the results
```

5.1.7. Composite trapezoidal rule.

Composite trapezoidal method for PDF of $N(0,1)$.

```
clear
clc

phi=@(x)((1/sqrt(2*pi))*exp(-x.^2/2)); %  PDF of nromal 0,1


a=0; % Left point of interval
b=1; % Right point of interval
d=0; % Assigning an initial value for integral to make the loop to start
n=0; % The number of sub-intervales in a partition
delta=0.0000001;

while(abs(d-normcdf(b))>delta) % Comparing to the real value
    n=n+1;
    h=(b-a)/n;
    c=0;
    for i=1:n-1
        c=c+2*(h/2)*phi(i*h+a); % Summing up the values on sub-intervals
    end

    ii(n,1)=1/2+c+(h/2)*phi(a)+(h/2)*phi(b); % Vlues for the integral
    d=ii(n,1);
end


csvwrite('Trap.dat',ii) % Writing the results in a file

ii % Display the values for the integral
```

5.1.8. Composite Simpson's rule. Composite Simpson's rule finding PDF of $N(0,1)$.

```
clear
clc

phi=@(x)((1/sqrt(2*pi))*exp(-x.^2/2)); %  PDF of nromal 0,1


a=0; % Left point of interval
b=1; % Right point of interval
d=0; % Assigning an initial value for integral
n=0; % The number of sub-intervales in a partition
delta=0.0000001;

m=0;
d=0;


while(abs(d-normcdf(b))>delta) % Copmare with the real value for integral
    m=m+1;
    n=2*m;
    h=(b-a)/n;
    c1=0;
    c2=0;
    for i=1:m
        c1=c1+(4*h/3)*phi((2*i-1)*h+a); % Summing up the odd indexes
    end

    for i=1:m-1
        c2=c2+(2*h/3)*phi((2*i)*h+a); % Summing up the even indexes
    end

    ii(m,1)=1/2+(h/3)*phi(a)+c1+c2+(h/3)*phi(b);
    % Summing up odd, even, first and last values

    d=ii(m,1);
end



csvwrite('Simp.dat',ii) % Write in an EXCEL file
ii % Display the results
```

5.2. Chapter 2 programs

5.2.1. Pricing with recombining binomial tree. This program has 5 parts. First the asset price on a binomial tree is found. Then, by backward induction European call, European put, American call and American put option prices are found, respectively.

```
clear
clc

%% Parameters
n=3; % The number of periods
s0=1; % Initial asset value
K=1; % Strike price
deltat=1; % Period length
sigma=0.3; % Volatility
r=0.045; % Interest rate

%% Finding necessary parameters
epsilon=sqrt((r+sigma^2/2)*deltat);
u=exp(epsilon);
nu=r-sigma^2/2;
pu=nu*sqrt(deltat)/(2*sqrt((r+sigma^2/2)))+0.5;
pd=1-pu;

%% Asset dynamic
for i=1:n
    for j=1:i
        S(j,i)=s0*u^(i-1-2*(j-1)); % Stock dynamic
    end
end

%% Options dynamic

for i=1:n
    for j=1:i
        EC(j,i)=max(S(j,i)-K,0); % EC= Call dynamic
    end
end

for i=1:n
    for j=1:i
        EP(j,i)=max(K-S(j,i),0); % EP=Put dynamic
    end
end

%% EU Call option pricing

PEC=zeros(size(EC));
PEC(:,end)=EC(:,end);

for i=n-1:-1:1
    for j=1:i
```

```
        PEC(j,i)=exp(-r*deltat)*(pu*PEC(j,i+1)+pd*PEC(j+1,i+1));
        %PEC= Price European call
    end
end


%% EU Put option price
PEP=zeros(size(EP));
PEP(:,end)=EP(:,end);

for i=n-1:-1:1
    for j=1:i
        PEP(j,i)=exp(-r*deltat)*(pu*PEP(j,i+1)+pd*PEP(j+1,i+1));
        %PEP= Price European put
    end
end


%% AM Call option pricing

PAC=zeros(size(EC));
PAC(:,end)=EC(:,end);

for i=n-1:-1:1
    for j=1:i
        PAC(j,i)=max(exp(-r*deltat)*(pu*PAC(j,i+1)+pd*PAC(j+1,i+1)),EC(j,i));
        %PAC= Price American call
    end
end

%% AM Call option pricing

PAP=zeros(size(EP));
PAP(:,end)=EP(:,end);

for i=n-1:-1:1
    for j=1:i
        PAP(j,i)=max(exp(-r*deltat)*(pu*PAP(j,i+1)+pd*PAP(j+1,i+1)),EP(j,i));
        %PAP= Price American put
    end
end


%% Display the value of the prices
[PEC(1,1);
PEP(1,1);
PAC(1,1);
PAP(1,1)]
```

5.2.2. Pricing with recombining trinomial tree. This program has 5 parts. First the asset price on a trinomial tree is found. Then, by backward induction European call, European put, American call and American put option prices are found, respectively.

```
clear
clc

%% Parameters
n=3;
S0=1;
K=1;
deltat=1;
sigma=0.3;
r=0.045;

%% Finding necessary parameters
epsilon=3*sqrt(deltat);
u=exp(epsilon);
nu=r-sigma^2/2;
pu=0.5*((sigma^2*deltat+nu^2*(deltat)^2)/epsilon^2+(nu*deltat)/epsilon);
pd=0.5*((sigma^2*deltat+nu^2*(deltat)^2)/epsilon^2-(nu*deltat)/epsilon);
pm=1-pu-pd;

%% Asset dynamic
for i=1:n
    for j=1:2*i-1
        S(j,i)=S0*u^(i-j); % Stock dynamic
    end
end

%% Options dynamic

for i=1:n
    for j=1:2*i-1
        EC(j,i)=max(S(j,i)-K,0); % EC= Call dynamic
    end
end

for i=1:n
    for j=1:2*i-1
        EP(j,i)=max(K-S(j,i),0); % EP=Put dynamic
    end
end

%% EU Call option pricing

PEC=zeros(size(EC));
PEC(:,end)=EC(:,end);

for i=n-1:-1:1
    for j=1:2*i-1
        PEC(j,i)=exp(-r*deltat)*(pu*PEC(j,i+1)+pm*PEC(j+1,i+1)+pd*PEC(j+2,i+1));
```

```
        %PEC= Price European call
    end
end


%% EU Put option price
PEP=zeros(size(EP));
PEP(:,end)=EP(:,end);

for i=n-1:-1:1
    for j=1:2*i-1
        PEP(j,i)=exp(-r*deltat)*(pu*PEP(j,i+1)+pm*PEP(j+1,i+1)+pd*PEP(j+2,i+1));
        %PEP= Price European put
    end
end


%% AM Call option pricing

PAC=zeros(size(EC));
PAC(:,end)=EC(:,end);

for i=n-1:-1:1
    for j=1:2*i-1
        PAC(j,i)=max(exp(-r*deltat)*(pu*PAC(j,i+1)+...
        pm*PAC(j+1,i+1)+pd*PAC(j+2,i+1)),EC(j,i));
        %PAC= Price American put
    end
end

%% AM Call option pricing

PAP=zeros(size(EP));
PAP(:,end)=EP(:,end);

for i=n-1:-1:1
    for j=1:2*i-1
        PAP(j,i)=max(exp(-r*deltat)*(pu*PAP(j,i+1)+...
        pm*PAP(j+1,i+1)+pd*PAP(j+2,i+1)),EP(j,i));
        %PAP= Price American put
    end
end

%% Display the value of the prices
[PEC(1,1);
PEP(1,1);
PAC(1,1);
PAP(1,1)]
```

5.3. Chapter 3 programs

5.3.1. Inversion & acceptance-rejection method. This is a program in Monte-Carlo method, Inversion & Acceptance-Rejection methods.

```
%% Parameters and functions

clear
clc
c=1.33;
n=100;
U=rand(n,1);
V=rand(n,1);

m=0; % Index of the acceptance

f=@(x)(2/(sqrt(2*pi))*exp(-x.^2/2));
g=@(x)(exp(-x));



%% Inversion method

EXP=log(1./(1-U));

EXP %Display exponential sample

%% Acceptance-rejection method

for i=1:n
    if (V(i,1)<=f(EXP(i,1))/c*g(EXP(i,1)))
        m=m+1;
        x(m,1)=EXP(i,1);
    end
end
x % Display the result

%% A norml sample
W=rand(length(x),1);
b=1-2*(W<=0.5);
y=b.*x;

y % Dsiplay the normal sample
```

5.3.2. Variance reduction method. This is a program in Monte-Carlo method, using variance reduction method to price a call option. The price is found by mean and mean using a control variable and then compare their variances.

```
clear
clc

%% Parameters
S0=1;
K=1;
sigma=0.3;
r=0.045;
N=100000;
RAN=normrnd(0,1,N,1);
T=2;


%% Asset price at time T
for i=1:N
    S(i,1)=S0*exp((r-sigma^2/2)*T+sigma*sqrt(T)*RAN(i,1));
end

%% Price and variances of call option

[exp(-r*T)*mean(max(S-K,0)) var(max(S-K,0))]

%% Price and variance of call option using variance reduction method

call=max(S-K);
b=((var(S+call)-var(S)-var(call))/2)/var(S);
Cb=max(S-K,0)-b*(S-S0*exp(r*T));

[exp(-r*T)*mean(Cb) var(Cb)]
```

5.3.3. Integral with simulation. Finding the integral by Monte-Carlo method, using a large number of simulations 100000000, that shows the lower accuracy w.r.t. trapezoidal and Simpson rules.

```
clear
clc
format long

f=@(x)(1/sqrt(2*pi)*exp(-x.^2/2));

M=1;
n=100000000; % The number of iterations

a=0;
b=1;

u=M*rand(n,1);
v=(b-a)*rand(n,1)+a;



A=1/2+M*(b-a)*sum((u<=f(v)))/n % The result

abs(normcdf(b)-A) % Compare with the real value
```

5.4. Chapter 4 programs

5.4.1. Explicit method. This is a program to use explicit method to find an approximation for $x' = 2x, x(0) = 1$.

```
clc
clear

f=@(x)(2*x);
n=2;

for i=3:5

    h=(1/n)^i;

    x(1,1)=1;

    for j=1:n^i

        x(j+1,1)=x(j,1)+h*f(x(j,1));

    end

    plot((0:h:1),x)
    hold on;


    x=zeros;

end

y=exp(2*(0:h:1));
plot((0:h:1),y)
hold on
```

5.4.2. Implicit method. This is a program to use implicit method to find an approximation for $x' = 2x, x(0) = 1$.

```
clc
clear

n=2;

for i=3:5

    h=(1/n)^i;

    x(1,1)=1;

    for j=1:n^i

        x(j+1,1)=((1+h)/(1-h))*x(j,1);

    end

    plot((0:h:1),x)
    hold on;


    x=zeros;

end

y=exp(2*(0:h:1));
plot((0:h:1),y)
hold on
```

5.4.3. Stochastic analysis and discretization. This is a program to simulate different process after discretization, including Brownian motion, Black-Scholes and Ornstein-Uhlenbeck process.

```
%% Parameters
clear
clc
T=1;
N=10;
n=1000;
h=T/n;
RAN=normrnd(0,1,N,n);

%%

for j=1:N
    B(j,1)=0;
    for i=1:n-1
        B(j,i+1)=B(j,i)+sqrt(h)*RAN(j,i+1);
    end
end


figure(1)
plot((0:h:1-h),B)
axis([-h T+h min(min(B))+h max(max(B))+h])

%% Black-Scholes model

S0=1;
mu=@(x)(0.05*x);
sigma=@(x)(0.1*x);

for j=1:N
    S(j,1)=S0;
    for i=1:n-1
        S(j,i+1)=S(j,i)+mu(S(j,i))*h+sigma(S(j,i))*sqrt(h)*RAN(j,i+1);
    end
end



figure(2)
plot((0:h:1-h),S)
axis([-h T+h min(min(S))+h max(max(S))+h])

%% Ornstein Uhlenbeck model

x0=0;
mu=@(x)((1-x));
sigma=@(x)(0.25);

for j=1:N
    x(j,1)=j-(N/2);
```

```
    for i=1:n-1
        x(j,i+1)=x(j,i)+mu(x(j,i))*h+sigma(x(j,i))*sqrt(h)*RAN(j,i+1);
    end
end



figure(3)
plot((0:h:1-h),x)
axis([-h T+h min(min(x))+h max(max(x))+h])
```

CHAPTER 6

Solutions to selected exercises

6.1. Chapter 1 selected solutions

6.1.1. Solution to exercise 23.

```
clear all

for i=1:5                                            % for loop to iterate all y's
    tol=0.00000001; a=0; b=1; delta=1; % declare all the parameters
while (delta>tol && counter<=10^10)                  % start iterate for bisection method
    % answer: c=(a+b)/2;
    f1=5*(a^2)-i;
    % answer: f2=5*(c^2)-i;
    f3=5*(b^2)-i;
    if f1*f2<=0
        b=c;
    else
        a=c;
    end
    % answer: delta = abs(5*(c^2)-i);

end                                                  % output results

end
```

6.1.2. Solution to exercise 24. Let us consider the function $f\left(x\right)=x^{2}-2$. Then we need to find the root of f. So we have to create the following iterative method

$$x_0=1.4, x_{k+1}=x_k-\frac{x_k^2-2}{2x_k}=x_k-\frac{1}{2}x_k+\frac{1}{x_k}=\frac{1}{2}x_k+\frac{1}{x_k}.$$

So to find the first iteration $x_1=\frac{1}{2}\left(1.4\right)+\frac{1}{1.4}\simeq 0.7+0.714=1.414$.

6.1.3. Solution to exercise 25.

(1) Using the Newton method we have to look at $x_{i+1}=x_i-\frac{f(x_i)}{f'(x_i)}$. So we have $x_{i+1}=x_i-\frac{(x_i-a)^L}{L(x_i-a)^{L-1}}=x_i-\frac{(x_i-a)}{L}=\frac{L-1}{L}x_i+\frac{a}{L}$.

(2) If the limit exists then x_i has to converge to that. If we denote it by x then we get:

$$x = \lim x_{i+1} = \lim \left(\frac{L-1}{L} x_i + \frac{a}{L} \right) = \frac{L-1}{L} x + \frac{a}{L}.$$

If we solve for x then we get $x = a$, so the limit is a.

(3) Let us prove the claim by induction. For $i = 1$, we only need to apply the Newton iteration: $x_1 = \frac{L-1}{L} x_0 + \frac{a}{L} = \frac{L-1}{L} (x_0 - a) + a$. Now let us consider the equality holds for i, i.e., $x_i = \left(\frac{L-1}{L}\right)^i (x_0 - a) + a$. Let us check it for $i+1$. Observe, from the Newton iteration that $x_{i+1} = \frac{L-1}{L} x_i + \frac{a}{L}$. Now replace x_i by $\left(\frac{L-1}{L}\right)^i (x_0 - a) + a$ to get $x_{i+1} = \left(\frac{L-1}{L}\right)^{i+1} (x_0 - a) + a$.

(4) The error term is simply $e_i = |x_i - 1| = \left| \left(\frac{3}{4}\right)^i \left(\frac{1}{4} - 1\right) + 1 - 1 \right| = \left|\frac{3}{4}\right|^{i+1}$. To make it smaller than δ, we need to look $\left|\frac{3}{4}\right|^{i+1} < 0.0001$, or $(i+1) \log\left(\frac{3}{4}\right) < \log(0.0001) = -4 \log(10)$ or $i > \frac{-4\log(10)}{\log\left(\frac{3}{4}\right)} - 1$. So at $i = \left\lfloor \frac{-4\log(10)}{\log\left(\frac{3}{4}\right)} \right\rfloor$, we reach the accuracy δ.

6.1.4. Solution to exercise 26. Take the values into $f(x, y)$ we have the first simplex: the

Best point (B): $f(2.4, 1.6) = -6.72$;

Good point (G): $f(3.6, 1.6) = -6.24$;

Worst point (W): $f(2.4, 2.4) = -6.24$.

(Note here the G and W are interchangeable, the following answer gives the situation when the above simplex is set)

Now we find the mid-point M of B and G is:

$$M = \frac{B+G}{2} = (3, 1.6)$$

Thus we have point the reflection R at:

$$R = 2M - W = (3.6, 0.8)$$

Observe $f(R) = f(3.6, 0.8) = -4.48 > -6.24 = f(3.6, 1.6) = f(B)$, so by Contraction we need to check another two points C_1 and C_2.

$$C_1 = \frac{M+R}{2} = (3.3, 1.2); \quad C_2 = \frac{W+M}{2} = (2.7, 2)$$

We have $f(C_1) = -6.03 > f(C_2) = -6.91$, and $f(C_2) < f(W)$ so replace W with C_2 and carry on the next iteration with a new complex: B(2.7,2); G(2.4,1.6); W(3.6,1.6). If the first chose by the student is: B(2.4,1.6); G(2.4,2.4) and W(3.6,1.6). Then we have: first mid-point M(2.4,2), then reflect to R(1.2,2.4) with $f(R) = -2.88$, so contract to $C_1 = (1.8, 2.2)$ and $C_2 = (3, 1.8)$ we have $f(C_1) = -5.28$ and $f(C_2) = -6.96$. Thus the new triangle is: B(3,1.8); G(2.4,1.6); W(2.4,2.4).

6.1.5. Solution to exercise 27. First we compute the steepest descent direction:

$$\nabla f(x, y) = (2x - y - 4, 2y - x - 1)$$

thus

$$\nabla f(\overrightarrow{x_0}) = \nabla f(3, 3) = (-1, 2)$$

$$||\nabla f(x_0)|| = \sqrt{1+4} = \sqrt{5}$$

$$(x_1, y_1) = (3, 3) - \gamma \frac{(-1, 2)}{\sqrt{5}} = (3 + \frac{\gamma}{\sqrt{5}}, 3 - \frac{2}{\sqrt{5}}\gamma)$$

(The following part is not necessary to learn)

If further we try to minimize the function in terms of γ

$$f(\overrightarrow{x_1}) = \frac{7}{5}\gamma^2 - \frac{5}{\sqrt{5}}\gamma - 6$$

with respect to γ, which yields:

$$\frac{df(\overrightarrow{x_1})}{d\gamma} = \frac{14}{5}\gamma - \frac{5}{\sqrt{5}} = 0$$
$$\Rightarrow \gamma_0 = \frac{5\sqrt{5}}{14}$$

Thus we can start the next iteration with point x_1 define as:

$$\overrightarrow{x_1} = \overrightarrow{x_0} - \frac{5\sqrt{5}}{14}\frac{\nabla f(x_0)}{\sqrt{5}} = (3,3) - \frac{5}{14}(-1,2) = (3\frac{5}{14}, 2\frac{2}{7}).$$

6.1.6. Solution to exercise 29. a) We have

$$\underbrace{\begin{bmatrix} 3 & 0 & 0 & 1 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 1 & 0 & 0 & 3 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}}_{\overrightarrow{x}^T} = \underbrace{\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}}_{b},$$

so it is $A\overrightarrow{x}^T = b$.
b)

$$A = \begin{bmatrix} 3 & 0 & 0 & 1 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 1 & 0 & 0 & 3 \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}}_{U} + \underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}}_{L} + \underbrace{\begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}}_{D}$$
$$= U + L + D.$$

c) Let

$$A\overrightarrow{x}^T = b \Rightarrow (U + L + D)\,\overrightarrow{x}^T$$
$$= b \Rightarrow D\overrightarrow{x}^T = b - (U + L)\,\overrightarrow{x}^T \Rightarrow \overrightarrow{x}^T = D^{-1}b - D^{-1}(U + L)\,\overrightarrow{x}^T.$$

So the function is, $g\left(\overrightarrow{x}^T\right) = D^{-1}b - D^{-1}(U+L)\overrightarrow{x}^T$. But we know $D^{-1} = \begin{bmatrix} \frac{1}{3} & 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 0 & \frac{1}{3} \end{bmatrix}$. Then the function is

$$g\left(\overrightarrow{x}^T\right) = D^{-1}b - D^{-1}(U+L)\overrightarrow{x}^T$$
$$= \begin{bmatrix} \frac{1}{3} & 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 0 & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} - \begin{bmatrix} \frac{1}{3} & 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 0 & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$
$$\begin{bmatrix} \frac{1}{3} \\ 0 \\ 0 \\ \frac{1}{3} \end{bmatrix} - \begin{bmatrix} \frac{x_4}{3} \\ 0 \\ 0 \\ \frac{x_1}{3} \end{bmatrix} = \begin{bmatrix} \frac{1-x_4}{3} \\ 0 \\ 0 \\ \frac{1-x_1}{3} \end{bmatrix}.$$

d) So if we want to find the fixed point of g we get

$$\begin{bmatrix} x_1^{k+1} \\ x_2^{k+1} \\ x_3^{k+1} \\ x_4^{k+1} \end{bmatrix} = \begin{bmatrix} \frac{1-x_4^k}{3} \\ 0 \\ 0 \\ \frac{1-x_1^k}{3} \end{bmatrix} \Rightarrow \begin{cases} x_1^{k+1} = \frac{1-x_4^k}{3} \\ x_2^{k+1} = 0 \\ x_3^{k+1} = 0 \\ x_4^{k+1} = \frac{1-x_1^k}{3} \end{cases}.$$

d) Finding x_4^k from last equality, and putting it inside the first equation we get

$$\begin{cases} x_1^{k+1} = \frac{2+x_1^{k-1}}{9} \\ x_2^{k+1} = 0 \\ x_3^{k+1} = 0 \\ x_4^{k+1} = \frac{2+x_4^{k-1}}{9} \end{cases}.$$

Since we assume the sequence is converging then we have

$$\begin{cases} x_1 = \frac{2+x_1}{9} \\ x_2 = 0 \\ x_3 = 0 \\ x_4 = \frac{2+x_4}{9} \end{cases},$$

which gives us $x_1 = x_4 = \frac{1}{4}, x_2 = x_3 = 0$.

6.1.7. Solution to exercise 32. We have

$$\begin{aligned}
\int_{-1}^{1} S(x)dx &= \int_{-1}^{0} S_1(x)dx + \int_{0}^{1} S_2(x)dx \\
&= \int_{-1}^{0} \left(-\frac{1}{4}(x+1)^3 + \frac{3}{4}x + \frac{5}{4}\right) dx \\
&+ \int_{0}^{1} \left(-\frac{1}{4}\left(1-x\right)^3 - \frac{3}{4}x + \frac{5}{4}\right) dx \\
&= \left(-\frac{1}{16}(x+1)^4 + \frac{3}{8}x^2 + \frac{5}{4}x\right)\Big|_{-1}^{0} \\
&+ \left(\frac{1}{16}(1-x)^4 - \frac{3}{8}x^2 + \frac{5}{4}x\right)\Big|_{0}^{1} \\
&= \left(-\frac{1}{16}\right) - \left(\frac{3}{8} - \frac{5}{4}\right) + \left(-\frac{3}{8} + \frac{5}{4}\right) - \left(\frac{1}{16}\right) \\
&= \frac{-1-6+20-6+20-1}{16} = \frac{13}{8}
\end{aligned}$$

Now, we use the Lagrangian polynomials.

$$L_0\left(x\right) = \frac{\left(x-0\right)\left(x-1\right)}{\left(-1-0\right)\left(-1-1\right)} = \frac{x\left(x-1\right)}{2}.$$

$$L_1\left(x\right) = \frac{\left(x-\left(-1\right)\right)\left(x-1\right)}{\left(0-\left(-1\right)\right)\left(0-1\right)} = 1 - x^2.$$

$$L_2\left(x\right) = \frac{\left(x-\left(-1\right)\right)\left(x-0\right)}{\left(-1-\left(-1\right)\right)\left(1-0\right)} = \frac{\left(x+1\right)x}{2}.$$

So the polynomial will be

$$\begin{aligned}
P_2\left(x\right) &= f\left(x_0\right)L_0\left(x\right) + f\left(x_1\right)L_1\left(x\right) + f\left(x_2\right)L_2\left(x\right) \\
&= \frac{x\left(x-1\right)}{4} + \left(1-x^2\right) + \frac{\left(x+1\right)x}{4} = 1 - \frac{x^2}{2}.
\end{aligned}$$

Therefore, the integral is $\int_{-1}^{1} P_2(x)\,dx = \int_{-1}^{1}\left(1-\frac{x}{2}^2\right)dx = x - \frac{x^3}{6}\Big|_{-1}^{1} = \frac{5}{3}$. The real integral is $\int_{-1}^{1}\frac{1}{1+x^2}dx = \arctan(x)|_{-1}^{1} = \frac{\pi}{4} - \left(-\frac{\pi}{4}\right) = \frac{\pi}{2} = 1.57$. The error for the spline is $\left|\frac{13}{8} - 1.57\right| \approx 0.05$ and $\left|\frac{5}{3} - 1.57\right| \approx 0.11$. So spline gives better answer.

6.1.8. Solution to exercise 33. a) Since we have four pairs of points, the minimum degree of the polynomial is 3. The general form of Lagrange polynomials is:

$$L_i(x) = \frac{\prod_{j\neq i}(x-x_j)}{\prod_{j\neq i}(x_i-x_j)}.$$

For $x_i = 1/1000, 1/8, 1, 8 \quad i = 1,2,3,4$ we have:

$$L_1(x) = \frac{x-1/8}{1/1000-1/8}\times\frac{x-1}{1/1000-1}\times\frac{x-8}{1/1000-8};$$

$$L_2(x) = \frac{x-1/1000}{1/8-1/1000}\times\frac{x-1}{1/8-1}\times\frac{x-8}{1/8-8};$$

$$L_3(x) = \frac{x-1/1000}{1-1/1000}\times\frac{x-1/8}{1-1/8}\times\frac{x-8}{1-8};$$

$$L_4(x) = \frac{x-1/1000}{8-1/1000}\times\frac{x-1/8}{8-1/8}\times\frac{x-1}{8-1}.$$

b) The general form of a fitted polynomial using Lagrange polynomials is:

$$f(x) = \sum_i y_i L_i(x),$$

and the fitted polynomial for $y_i = 1/10, 1/2, 1, 2 \quad i = 1,2,3,4$ is:

$$f(x) = \frac{1}{10}L_1(x) + \frac{1}{2}L_2(x) + L_3(x) + 2L_4(x).$$

c) The above polynomial gives an approximation of $\sqrt[3]{2}$ of -2.19099453. Compare with $\sqrt[3]{2} \approx 1.25992104$, the approximation is bad and with an error of 3.45091558.

d) In the lecture, the error is give by:

$$E_n(x) = \frac{\prod_i(x-x_i)f^{(n+1)}(c)}{(n+1)!}.$$

where $c \in [a, b]$. In our case, $n = 3$, $x_i = 1/1000, 1/8, 1, 8 \quad i = 1, 2, 3, 4$, $f(x) = \sqrt[3]{x}$ and $[a, b] = [1/1000, 8]$. First we calculate:

$$f^{(n+1)}(x) = f^{(3+1)}(x) = (\sqrt[3]{x})^{(4)} = -\frac{80}{81}x^{-11/3},$$

and this is a monotone increasing function in $x \in [1/1000, 8]$. Also we have:

$$\frac{\prod_i (2 - x_i)}{4!} = -0.93703125,$$

so we get the error is estimated within the range:

$$-0.93703125 \times [f^{(4)}(8), f^{(4)}(0.001)] = [4.51886212e{-}04, 9.25462963e{+}10].$$

Thus, we find that the actual error 3.45091558 is within the estimated error arrange.

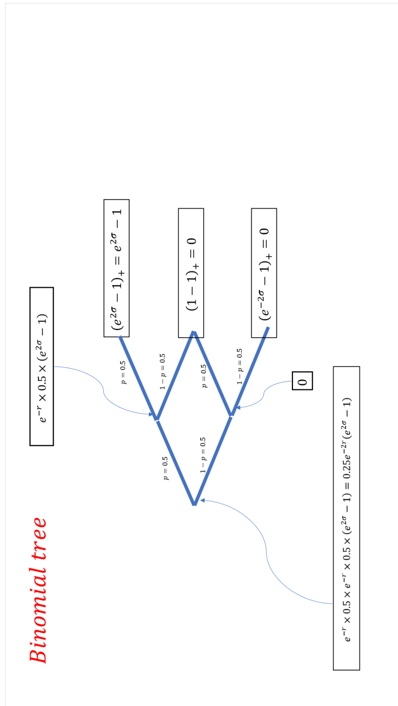
6.2. Chapter 2 selected solutions

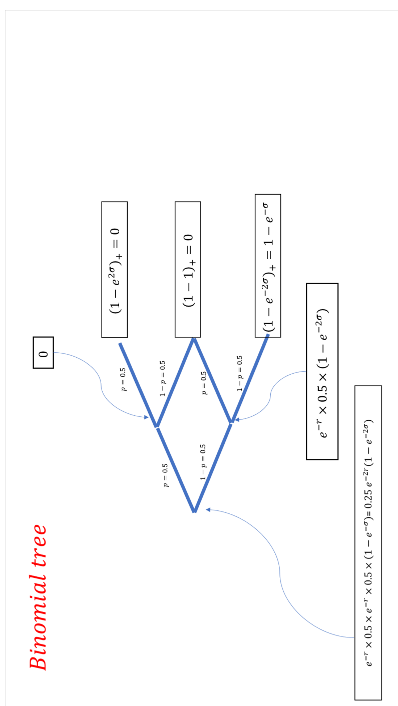
6.2.1. Solution to exercise 43. a) For the binomial tree the parameter ϵ is found by

$$\epsilon = \sqrt{\sigma^2\delta t + \left(\overbrace{\left(r - \frac{\sigma^2}{2}\right)\delta t}^{=0}\right)^2} = \sqrt{\sigma^2 \times 1} = \sigma,$$

The probability though is given by

$$p = \frac{\overbrace{\left(r - \frac{\sigma^2}{2}\right)\delta t}^{=0}}{2\epsilon} + \frac{1}{2} = \frac{1}{2}.$$





b) For the trinomial tree the parameter ϵ is found by

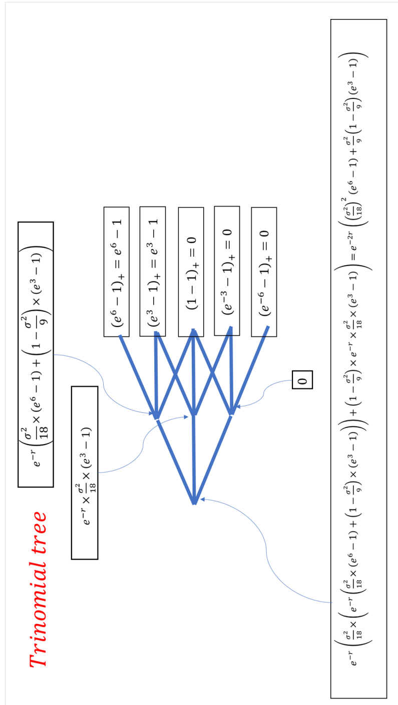
$$\epsilon = 3\sqrt{\delta t} = 3.$$

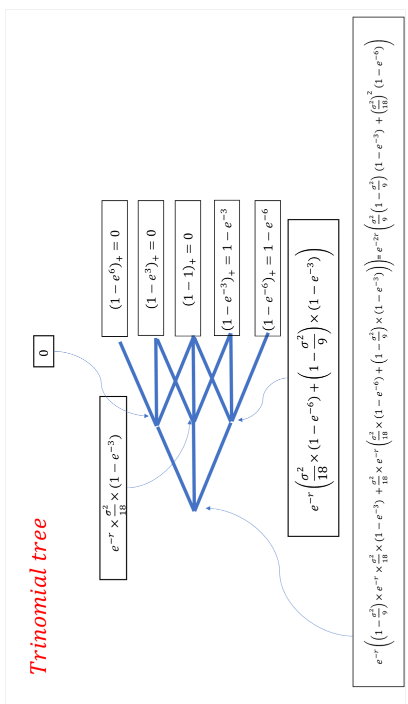
Since $v = \left(r - \frac{\sigma^2}{2}\right) = 0$The probability though is given by

$$p_u = \frac{1}{2}\left(\frac{\sigma^2 \delta t + v^2(\delta t)^2}{\epsilon^2} + \frac{v\delta t}{\epsilon}\right) = \frac{\sigma^2}{18},$$

$$p_m = 1 - \left(\frac{\sigma^2 \delta t + v^2(\delta t)^2}{\epsilon^2}\right) = 1 - \frac{\sigma^2}{9},$$

$$p_d = \frac{1}{2}\left(\frac{\sigma^2 \delta t + v^2(\delta t)^2}{\epsilon^2} - \frac{v\delta t}{\epsilon}\right) = \frac{\sigma^2}{18}.$$


Trinomial tree
$(e^6-1)_+ = e^6-1$
$(e^3-1)_+ = e^3-1$
$(1-1)_+ = 0$
$(e^{-3}-1)_+ = 0$
$(e^{-6}-1)_+ = 0$
$e^{-r}\left(\frac{\sigma^2}{18}\times(e^6-1)+\left(1-\frac{\sigma^2}{9}\right)\times(e^3-1)\right)$
$e^{-r}\times\frac{\sigma^2}{18}\times(e^3-1)$
0


Trinomial tree
$(1-e^6)_+ = 0$
$(1-e^3)_+ = 0$
$(1-1)_+ = 0$
$(1-e^{-3})_+ = 1-e^{-3}$
$(1-e^{-6})_+ = 1-e^{-6}$
0
$e^{-r}\times\frac{\sigma^2}{18}\times(1-e^{-3})$
$e^{-r}\left(\frac{\sigma^2}{18}\times(1-e^{-6})+\left(1-\frac{\sigma^2}{9}\right)\times(1-e^{-3})\right)$

6.3. Chapter 3 selected solutions

6.3.1. Solution to exercise 55.

(1) Let $Y = |X|^{\frac{1}{r}}$. First we want to find F_Y and its inverse F_Y^{-1}, in terms of F_X and F_X^{-1}, respectively. It is clear that since $Y \geq 0$, $F_Y(x) = 0$, for $x \leq 0$. Noe consider $x > 0$, we have:

$$\begin{aligned} F_Y(x) &= P(Y \leq x) = P\left(|X|^{\frac{1}{r}} \leq x\right) \\ &= P\left(|X| \leq x^r\right) = P\left(-x^r \leq X \leq x^r\right) \\ &= \int_{-x^r}^{x^r} f_X(s)ds = 2\int_0^{x^r} f_X(s)ds \\ &= 2\left(\int_{-\infty}^{x^r} f_X(s)ds - \underbrace{\int_{-\infty}^{0} f_X(s)ds}_{1/2}\right) = 2F_X\left(x^r\right) - 1. \end{aligned}$$

So then, to find F_Y^{-1}, we have to first for some a, get $a = F_Y\left(x\right) = 2F_X\left(x^r\right) - 1$. Consequently, we get $x = \left|F_X^{-1}\left(\frac{a+1}{2}\right)\right|^{\frac{1}{r}}$, and therefore, $F_Y^{-1}\left(a\right) = \left|F_X^{-1}\left(\frac{a+1}{2}\right)\right|^{\frac{1}{r}}$.

(2) Now let's assume $f_X(x) = \frac{b}{2}e^{-b|x|}$. Then, we have to find $F_X\left(x\right) = \int_{-\infty}^{x} \frac{b}{2}e^{-b|s|}ds$. First, let us assume that $x \leq 0$, $F_X\left(x\right) = \int_{-\infty}^{x} \frac{b}{2}e^{-b(-s)}ds = \frac{b}{2b}e^{bs}|_{-\infty}^{x} = \frac{1}{2}e^{bx}$.

Second, assume $x > 0$, then $F_X\left(x\right) = \int_{-\infty}^{0} \frac{b}{2}e^{b|s|}ds + \int_0^x \frac{b}{2}e^{-bs}ds = \frac{1}{2} + \frac{1}{2}\left(-e^{-bx} + 1\right) = 1 - \frac{1}{2}e^{-bx}$. So in total we have

$$F_X\left(x\right) = \begin{cases} \frac{1}{2}e^{bx}, & x \leq 0 \\ 1 - \frac{1}{2}e^{-bx}, & x > 0 \end{cases}.$$

Now let us see what is F_X^{-1}. For that we have to consider two cases $0 \leq a \leq \frac{1}{2}$ and $\frac{1}{2} \leq a \leq 1$. In the first case we have $a = F_X\left(x\right) = \frac{1}{2}e^{bx}$, so we get $x = \frac{\log(2a)}{b}$. In the second case we have $a = F_X\left(x\right) = 1 - \frac{1}{2}e^{-bx}$, and as a result we get $x = \frac{\log(2(1-a))}{-b}$.

In total we have

$$F_X^{-1}(a) = \begin{cases} \frac{\log(2a)}{b} & 0 \le a \le \frac{1}{2} \\ \frac{\log(2(1-a))}{-b} & \frac{1}{2} \le a \le 1 \end{cases}.$$

Based on discussions we had in part 1, we know

$$F_Y^{-1}(a) = \left| F_X^{-1}\left(\frac{a+1}{2}\right)\right|^{\frac{1}{r}} = \begin{cases} \left|\frac{\log(a+1)}{b}\right|^{\frac{1}{r}} & -1 \le a \le 0 \\ \left|\frac{\log(1-a)}{b}\right|^{\frac{1}{r}} & 0 \le a \le 1 \end{cases}$$

$$= \left|\frac{\log(1-a)}{b}\right|^{\frac{1}{r}} \text{(since always } 0 \le a \le 1)$$

(3) Now considering $b = 1$, and $r = 2$, with a $U(0,1)$ sample $u_1 = 1/4, u_2 = 1/2$ and $u_3 = 3/4$. If we want to generate a sample for Y and compute the empirical mean of Y we have:

$$y_1 = F_Y^{-1}\left(\frac{1}{4}\right) = \sqrt{\left|\log\left(\frac{3}{4}\right)\right|} = 0.35346674045,$$

$$y_2 = F_Y^{-1}\left(\frac{1}{2}\right) = \sqrt{\left|\log\left(\frac{1}{2}\right)\right|} = 0.54866200493,$$

$$y_1 = F_Y^{-1}\left(\frac{3}{4}\right) = \sqrt{\left|\log\left(\frac{1}{4}\right)\right|} = 0.77592524854.$$

The average mean is

$$\frac{0.77592524854 + 0.54866200493 + 0.35346674045}{3} = 0.55935133131.$$

6.3.2. Solution to exercise 56. 1. Since $-|x| - \frac{x^2}{2} \le -|x|$, and since H is an increasing function then we have $cf(x) = H\left(-|x| - \frac{x^2}{2}\right) \le H(-|x|) = g(x)$. So as a result, $\frac{f(x)}{g(x)} \le \frac{1}{c}$.

2. $c = \int_{-\infty}^{\infty} H\left(-|x| - \frac{x^2}{2}\right)dx = \int_{-\infty}^{\infty} \frac{1}{2} e^{\left(-|x|-\frac{x^2}{2}\right)} dx$. But since $\frac{1}{2}e^{\left(-|x|-\frac{x^2}{2}\right)}$ is a symmetric function we have $\int_{-\infty}^{\infty} \frac{1}{2}e^{\left(-|x|-\frac{x^2}{2}\right)}dx = 2\int_0^{\infty} \frac{1}{2}e^{\left(-|x|-\frac{x^2}{2}\right)}dx$. Therefore, we have,

$$
\begin{aligned}
c &= \frac{2}{2}\int_0^\infty e^{-\frac{1}{2}\left(x^2+2|x|+1\right)+\frac{1}{2}}dx \\
&= e^{\frac{1}{2}}\int_0^\infty e^{-\frac{1}{2}(x+1)^2}dx = e^{\frac{1}{2}}\int_1^\infty e^{-\frac{1}{2}x^2}dx \\
&= \sqrt{2\pi}e^{\frac{1}{2}}\frac{1}{\sqrt{2\pi}}\int_1^\infty e^{-\frac{1}{2}x^2}dx = \sqrt{2\pi}e^{\frac{1}{2}}\left(1-\Phi\left(1\right)\right),
\end{aligned}
$$

where Φ is the CDF of $N(0,1)$. From exercise 56 we know that for PDF, g the inverse of the CDF is given as

$$
G^{-1}\left(a\right) = \begin{cases} \log\left(2a\right) & 0 \le a \le \frac{1}{2} \\ \log\left(2\left(1-a\right)\right) & \frac{1}{2} \le a \le 1 \end{cases} = \log\left(1-|2a-1|\right).
$$

Now, we use acceptance-rejection and the inversion method combined, and we have the following algorithm:

(1) Generate two independent $U\left(0,1\right)$ sample $u_1,...,u_n$ and $v_1,...,v_n$.

(2) Generate a sample for PDF, g, as follows $y_1 = \log\left(1-|2u_1-1|\right),...,y_n = \log\left(1-|2u_n-1|\right)$.

(3) For every $1 \le i \le n$, if

$$
\begin{aligned}
v_i \le \frac{f\left(y_i\right)}{\left(1/c\right)g\left(y_i\right)} &= \frac{\sqrt{2\pi}e^{\frac{1}{2}}\left(1-\Phi\left(1\right)\right)e^{-\left(|\log(1-|2u_i-1|)|+\frac{\left(\log\left(1-|2u_i-1|\right)\right)^2}{2}\right)}}{\left(\frac{1}{2}e^{-|\log(1-|2u_i-1|)|}\right)} \\
= \frac{\sqrt{2\pi}e^{\frac{1}{2}}\left(1-\Phi\left(1\right)\right)e^{-\left(|\log(1-|2u_i-1|)|+\frac{\left(\log\left(1-|2u_i-1|\right)\right)^2}{2}\right)}}{\left(\frac{1}{2}e^{-|\log(1-|2u_i-1|)|}\right)} \\
= 2\sqrt{2\pi}e^{\frac{1}{2}}\left(1-\Phi\left(1\right)\right)e^{-\left(\frac{\left(\log\left(1-|2u_i-1|\right)\right)^2}{2}\right)},
\end{aligned}
$$

then accept $\log\left(1-|2u_i-1|\right)$ as a sample, otherwise reject.

6.3.3. Solution to exercise 57. a) First of all observe that $f_{n,0}(x) = b_{n,0}x^n$, so for $0 \le x \le 1$, we have $F_{n,0}\left(x\right) = \int_0^x b_{n,0}s^n ds = \frac{b_{n,0}}{n+1}x^{n+1}$. On the other hand, we have $b_{n,0} = \left(\int_0^1 x^n dx\right)^{-1} = \left(\frac{1}{n+1}\right)^{-1} = n+1$. So, we get $F_{n,0}\left(x\right) = \frac{n+1}{n+1}x^{n+1} = x^{n+1}$ for $0 \le x \le 1$.

b) First we need to find the inverse of $F_{n,0}$. It is easy to see that $y = F_{n,0}(x) = x^{n+1}$ then $F_{n,0}^{-1}(y) = y^{\frac{1}{n+1}}$. Let's $u_1, ..., u_N$ be a $U(0,1)$ sample, then by inversion method a sample whose CDF is $F_{n,0}$ is given by $u_1^{\frac{1}{n+1}}, ..., u_N^{\frac{1}{n+1}}$.

c) Let us consider the fraction $\frac{f_{n.m}(x)}{f_{n,0}(x)} = \frac{b_{n,m}x^n(1-x)^m}{b_{n,0}x^n} = \frac{b_{n,m}(1-x)^m}{b_{n,0}}$. Note that since $0 \leq x \leq 1$, then $(1-x)^m \leq 1$ and the maximum of it is equal to 1. As a result $\frac{b_{n,m}}{b_{n,0}}$ is the smallest number c that $\frac{f_{n.m}(x)}{f_{n,0}(x)} \leq c$. So let us take $c = \frac{b_{n,m}}{b_{n,0}}$. So to design an acceptance-rejection algorithm we need to consider the fraction $\frac{f(x)}{cg(x)} = \frac{f_{n.m}(x)}{\frac{b_{n,m}}{b_{n,0}}f_{n,0}(x)} = \frac{b_{n,m}x^n(1-x)^m}{\frac{b_{n,m}}{b_{n,0}}b_{n,0}x^n} = (1-x)^m$. The algorithm is as follows

1) Generate two independent $U(0,1)$ samples $u_1, ..., u_N$ and $v_1, ..., v_N$.

2) By inversion method generate a sample for PDF $f_{n,0}$ i.e., $y_1 = u_1^{\frac{1}{n+1}}, ..., y_N = u_N^{\frac{1}{n+1}}$.

3) For any $1 \leq i \leq N$, if $v_i \leq \frac{f(y_i)}{cg(y_i)} = \left(1 - u_i^{\frac{1}{n+1}}\right)^m$ then accept $u_1^{\frac{1}{n+1}}$ as a sample for PDF $f_{n,m}$, otherwise, reject.

d) We have to compare the following

$$0.01 \leq \left(1 - (0.64)^{\frac{1}{2}}\right)^2 = (1-0.8)^2 = 0.04.$$

$$0.49 \leq \left(1 - (0.04)^{\frac{1}{2}}\right)^2 = (1-0.2)^2 = 0.64.$$

$$0.04 \leq \left(1 - (0.81)^{\frac{1}{2}}\right)^2 = (1-0.9)^2 = 0.01.$$

Only sample one and two are accepted so a sample is 0.2 and 0.8.

6.3.4. Solution to exercise 58. a) By changing variable we have

$$G_X(x) = L\int_0^x \frac{s^k}{1+s^{2k+2}}dx = L\int_0^x \frac{s^k}{1+s^{2k+2}}ds$$
$$\underbrace{=}_{s^{k+1}\leftrightarrow S} \frac{\frac{2k+2}{\pi}}{k+1}\int_0^{x^{k+1}} \frac{dS}{1+S^2} = \frac{2}{\pi}\tan^{-1}\left(x^{k+1}\right).$$

Note that

$$\int_0^\infty \frac{x^k}{1+x^{2k+2}}dx = \frac{1}{k+1}\int_0^\infty \frac{1}{1+x^{2k+2}}d\left(x^{k+1}\right)$$
$$= \frac{1}{k+1}\int_0^\infty \frac{1}{1+S^2}dS = \frac{1}{k+1}\tan^{-1}(\infty)$$
$$= \frac{\pi}{2k+2}.$$

So, if we take $y = G_X(x)$ then $x = \left(\tan\left(\frac{\pi}{2}y\right)\right)^{\frac{1}{k+1}}$. So, we have $G_X^{-1}(y) = \left(\tan\left(\frac{\pi}{2}y\right)\right)^{\frac{1}{k+1}}$. Now, for a given $U(0,1)$ sample $u_1, ..., u_n$, one can generate a sample with CDF G_X as follows

$$\left(\tan\left(\frac{\pi}{2}u_1\right)\right)^{\frac{1}{k+1}}, ..., \left(\tan\left(\frac{\pi}{2}u_n\right)\right)^{\frac{1}{k+1}}.$$

b) By assumption we get $P_{2k+1}(x) + x^{2k+2} \geq 1 + x^{2k+2}$. Based on this we get

$$\frac{f(x)}{g(x)} = \frac{\frac{ax^k}{P_{2k+1}(x)+x^{2k+2}}}{L\frac{x^k}{1+x^{2k+2}}} = \frac{a}{L}\frac{1+x^{2k+2}}{P_{2k+1}(x)+x^{2k+2}}$$
$$\leq \frac{a}{L}.$$

But on the other hand if we send x to ∞ then we have $\frac{a}{L}\frac{1+x^{2k+2}}{P_{2k+1}(x)+x^{2k+2}} \to \frac{a}{L}$. This means the supremum of $\frac{f(x)}{g(x)}$ is $\frac{a}{L}$.

c) In order to design an acceptance-rejection algorithm we need to look at the fraction

$$\frac{f(x)}{cg(x)} = \frac{a\left(1+x^{2k+2}\right)}{\frac{a}{L}\left(L\left(P_{2k+1}(x)+x^{2k+2}\right)\right)} = \frac{1+x^{2k+2}}{P_{2k+1}(x)+x^{2k+2}}.$$

Now we can introduce the following algorithm

1) Generate two independent $U(0,1)$ samples $u_1, ..., u_n$ and $v_1, ..., v_n$.

2) Generate a sample for CDF G_X as

$$\left(\tan\left(\frac{\pi}{2}u_1\right)\right)^{\frac{1}{k+1}}, ..., \left(\tan\left(\frac{\pi}{2}u_n\right)\right)^{\frac{1}{k+1}}.$$

3) If for any $1 \leq i \leq n$,

$$v_i \leq \frac{1+\left(\left(\tan\left(\frac{\pi}{2}u_i\right)\right)^{\frac{1}{k+1}}\right)^{2k+2}}{P_{2k+1}\left(\left(\tan\left(\frac{\pi}{2}u_i\right)\right)^{\frac{1}{k+1}}\right)+\left(\left(\tan\left(\frac{\pi}{2}u_i\right)\right)^{\frac{1}{k+1}}\right)^{2k+2}} = \frac{1+\left(\tan\left(\frac{\pi}{2}u_i\right)\right)^2}{P_{2k+1}\left(\left(\tan\left(\frac{\pi}{2}u_i\right)\right)^{\frac{1}{k+1}}\right)+\left(\tan\left(\frac{\pi}{2}u_i\right)\right)^2}$$

then accept $\left(\tan\left(\frac{\pi}{2}u_i\right)\right)^{\frac{1}{k+1}}$ as a sample otherwise reject.

d) For this part just let's have a hint. Since $\frac{f(x)}{cg(x)}$ is between 0 and 1, for v_1, v_3, v_5 we accept and for the other two we reject.

6.3.5. Solution to exercise 60. First we want to find the standard deviation $\sigma(Y(b))$, in terms of standard deviations $\sigma(Y)$, $\sigma(X)$ and number b. For that observe that

$$\begin{aligned} Var(Y(b)) &= E\left(Y(b)^2\right) - E(Y(b))^2 \\ &= E\left((Y-bX)^2\right) - \left(E\left(Y-bX\right)\right)^2 \\ &= E\left(Y^2\right) - 2bE\left(XY\right) + b^2E\left(X^2\right) \\ &\qquad - E\left(Y\right)^2 + 2bE\left(X\right)E\left(Y\right) - b^2E\left(X\right)^2 \\ &= \sigma_Y^2 + b^2\sigma_X^2 - 2b\left(E\left(XY\right) - E\left(X\right)E\left(Y\right)\right). \end{aligned}$$

So now we have to find $E\left(XY\right)$ and $E\left(X\right)E\left(Y\right)$. First, given the fact that $f_X\left(x\right) = f_X\left(-x\right)$ we have

$$\begin{aligned} E\left(X\right) &= \int_{-\infty}^{\infty} xf_X\left(x\right)dx \\ &= \int_{-\infty}^{0} xf_X\left(x\right)dx + \int_{0}^{\infty} xf_X\left(x\right)dx \\ &= -\int_{0}^{\infty} xf_X\left(x\right)dx + \int_{0}^{\infty} xf_X\left(x\right)dx = 0. \end{aligned}$$

Second, again given the fact that $f_X\left(x\right) = f_X\left(-x\right)$ we have

$$\begin{aligned} E\left(XY\right) &= E\left(X \times X1_{\{X\geq 0\}}\right) = E\left(X^2 1_{\{X\geq 0\}}\right) \\ &= \int_{0}^{\infty} x^2 f_X\left(x\right)dx = \frac{1}{2}\int_{-\infty}^{\infty} x^2 f_X\left(x\right)dx \\ &= \frac{1}{2}E\left(X^2\right) = \frac{1}{2}\left(E\left(X^2\right) - E\left(X\right)^2\right) = \frac{1}{2}\sigma_X^2. \end{aligned}$$

Now if we implement the last two equality in above we get

$$\begin{aligned} Var(Y(b)) &= E\left(Y(b)^2\right) - E(Y(b))^2 \\ &= E\left((Y-bX)^2\right) - \left(E\left(Y-bX\right)\right)^2 \\ &= E\left(Y^2\right) - 2bE\left(XY\right) + b^2E\left(X^2\right) \\ &\qquad - E\left(Y\right)^2 + 2bE\left(X\right)E\left(Y\right) - b^2E\left(X\right)^2 \\ &= \sigma_Y^2 + b^2\sigma_X^2 - b\sigma_X^2 = \sigma_Y^2 + b\left(b-1\right)\sigma_X^2. \end{aligned}$$

Now we can find the minimum of $Var\left(Y\left(b\right)\right)$ by minimizing $g(b) = b\left(b-1\right)$. The minimum is easily happening at $2b^{*}-1 = g'(b^{*}) = 0$, or after solving at $b^{*} = \frac{1}{2}$. So if we set $b^{*} = \frac{1}{2}$, then the minimum variance is given when we use the control variable

$$\begin{aligned} Y\left(\frac{1}{2}\right) &= Y - \frac{1}{2}X = X1_{\{X\geq 0\}} - \frac{1}{2}X \\ &= \begin{cases} X - \frac{1}{2}X = \frac{1}{2}X & X \geq 0 \\ -\frac{1}{2}X & X \leq 0 \end{cases} = \frac{1}{2}\left|X\right|. \end{aligned}$$

This means for a sample $x_1, ..., x_n$. The sample mean for Y is $\frac{x_1 1_{\{x_1\geq 0\}}+\cdots+x_n 1_{\{x_n\geq 0\}}}{n}$. Using the control variable the average is more efficiently can be estimated by $\frac{|x_1|+\cdots+|x_n|}{2n}$.

6.3.6. Solution to exercise (59)**.** a) We have

$$\begin{aligned} E\left(h\left(X\right)\right) &= \int_{-\infty}^{\infty} h\left(x\right) f_X\left(x\right) dx = \int_{-\infty}^{0} h\left(x\right) f_X\left(x\right) dx + \int_{0}^{\infty} h\left(x\right) f_X\left(x\right) dx \\ &= \int_{0}^{\infty} \underbrace{h\left(-x\right)}_{=-h(x)} \underbrace{f_X\left(-x\right)}_{=f_X(x)} dx + \int_{0}^{\infty} h\left(x\right) f_X\left(x\right) dx \\ &= -\int_{0}^{\infty} h\left(x\right) f_X\left(x\right) dx + \int_{0}^{\infty} h\left(x\right) f_X\left(x\right) dx = 0 \end{aligned}$$

b) Let us find the variance of Y_b. But first since the function $h(x) = x$ is odd, then based on part a) $E\left(X\right) = E\left(h\left(X\right)\right) = 0$. Now we have

$$\begin{aligned} Var\left(Y_b\right) &= Var\left(Y - b(X - E(X))\right) = Var\left(Y - bX\right) \\ &= Var\left(Y\right) + b^2 Var\left(X\right) - 2bcov\left(Y, X\right) \\ &= Var\left(Y\right) + b^2 Var\left(X\right) - 2b\left(E\left(XY\right) - \underbrace{E\left(X\right)}_{=0} E\left(Y\right)\right). \end{aligned}$$

Observe that $E\left(XY\right) = E\left(XY\right) = E\left(Xg\left(X\right)\right)$. Since g is even then $x \mapsto xg\left(x\right)$ is an odd function and using part a) we get $E\left(Xg\left(X\right)\right) = 0$. So we get that $Var\left(Y_b\right) = Var\left(Y\right) + b^2 Var\left(X\right)$. As a result, the minimum is taken at $b = 0$, and variance cannot be reduced.

c) First we have

$$E\left(Y_b^2\right) = E\left(Y^2\right) - 2bE\left(XY\right) + b^2E\left(X^2\right).$$

Since $E\left(Y_b\right) = E\left(Y\right)$ then we get

$$Var\left(Y_b\right) = Var\left(Y\right) - 2bE\left(Xg\left(X\right)\right) + b^2E\left(X^2\right).$$

As a result the optimal b can be found by equalizing derivative to zero: $b^* = \frac{E(Xg(X))}{E(X^2)}$.

d) Now since g is odd we get $-Xg\left(-X\right) = XgX$ and for the optimal b^* we have

$$\begin{aligned}
b^* = \frac{E\left(Xg\left(X\right)\right)}{E\left(X^2\right)} &= \frac{E\left(Xg\left(X\right)1_{\{X\geq 0\}}\right)}{E\left(X^2\right)} + \frac{E\left(Xg\left(X\right)1_{\{X<0\}}\right)}{E\left(X^2\right)} \\
&= \frac{E\left(Xg\left(X\right)1_{\{X\geq 0\}}\right)}{E\left(X^2\right)} + \frac{E\left(-Xg\left(-X\right)1_{\{X<0\}}\right)}{E\left(X^2\right)} \\
&= \frac{E\left(\left|X\right|g\left(\left|X\right|\right)1_{\{X\geq 0\}}\right)}{E\left(X^2\right)} + \frac{E\left(\left|X\right|g\left(\left|X\right|\right)1_{\{X<0\}}\right)}{E\left(X^2\right)} \\
&= \frac{E\left(\left|X\right|g\left(\left|X\right|\right)\right)}{E\left(X^2\right)} \leq \frac{\frac{1}{2}E\left(\left|X\right|^2\right)}{E\left(X^2\right)} = \frac{1}{2}.
\end{aligned}$$

6.3.7. Solution to exercise 62.

```
clc;clear all;
% Declare f(x)
f=@(x) 1./(1+0.1.*x^2);
% we know f(x) has its maximum at x = 0
M = 1;
% simulation trails
trails = 10^6;
% record the number of dots drop inside the function
number_inside = 0;
% bounds
a = -10; b = 10;
for i = 1:trails
    U = unifrnd(0,M);
    V = unifrnd(a,b);
    if U <= f(V)
        number_inside =  number_inside + 1;
    end
end
```

6.4. Chapter 4 selected solutions

6.4.1. Solution to exercise 82. a) For some $u > v$ one can have:

$$|f(t,u) - f(x,v)| = 2x^{-4}|u-v| \leq 2|u-v|,$$

so Lipschitz condition holds and we have $L = 2$.
b) For some $u > v$ it is easy to see:

$$\begin{aligned} |f(t,u) - f(t,v)| &= 2(1+e^{-|t|})\left|\frac{u}{1+u^2} - \frac{v}{1+v^2}\right| \\ &= 2(1+e^{-|t|})\frac{|1-uv|}{(1+u^2)(1+v^2)}|u-v| \end{aligned}$$

Also for $t \in (-\infty, \infty)$, $(1+e^{-|t|}) \leq 2$ and

$$\frac{|1-uv|}{(1+u^2)(1+v^2)} \leq \frac{1+|uv|}{(1+u^2)(1+v^2)} \leq 1 = 2(1+e^{-|t|})\frac{|1-uv|}{(1+u^2)(1+v^2)}|u-v|$$

thus we have:

$$|f(t,u) - f(t,v)| \leq 2 \times 2 \times 1 \times |u-v|,$$

so Lipschitz condition holds and we have $L = 4$.

6.4.2. Solution to exercise 83. a) A general form is given as

$$x' = f(t,x)$$

Therefore,

$$x' = -2t^2 x$$

where

$$f(t,x) = -2t^2 x$$

is a general form for our ODE.

b) First of all, we have to show the function f can be controlled over the first argument given that $x = x_0$, i.e.

$$\exists K \quad |f(t,x_0)| \leq K, \qquad \forall t \in [0,T]$$

So, if $x = x_0 \;\Rightarrow\; f(t,x_0) = f(t,1) = -2t^2$

$$t \in [0,T] \;\Rightarrow\; |f(t,x_0)| \leq 2T^2$$

Therefore, $K = 2T^2$ is an appropriate choice.

Secondly, we have to find a Lipschitz constant L_f over the second argument while we are fixing the first argument, i.e.

$$\exists C \quad |f(t,u) - f(t,v)| \leq L_f \, |u - v|$$
$$x_0 - C \leq u, v \leq x_0 + C$$
$$C \geq \frac{K}{L}\left(e^{L(t_M - t_0)} - 1\right)$$

So, it is clear that

$$\begin{aligned}\forall u, v, |f(t,u) - f(t,v)| &= \left|-2t^2 u + 2t^2 v\right| \\ &= \left|2t^2\right| |u - v| \\ &\leq 2T^2 \, |u - v|\end{aligned}$$

So $L = 2T^2$ might be a good candidate, however, we have to also check whether there is a C that satisfies the condition above.

$$K = 2T^2, L = 2T^2 \quad \Rightarrow \quad C \geq e^{2T^3} - 1$$

If we take $C = e^{2T^3} - 1$, then everything works out since the Lipschitz inequality holds for any u, v including

$$1 - (e^{2T^3} - 1) \leq u, v \leq (e^{2T^3} - 1) + 1$$

Therefore, according to Picard's theorem, there exists a solution on $[0, T]$ which is bounded as

$$|x - 1| \leq e^{2T^3} - 1, \quad 0 \leq t \leq T$$

c) Now if

$$t_i = \frac{iT}{n}, \quad 0 \leq i \leq n$$

be a partition for $[0, T]$, then, Euler explicit is given by:

$$\begin{cases} x_{i+1} = x_i + hf(t_i, x_i) & \qquad i = 1, ..., n-1 \\ x(0) = x_0 & \end{cases}$$

$$\begin{cases} x_{i+1} = x_i - \dfrac{B}{n}\left(2\left(\dfrac{iT}{n}\right)^2\right) x_i & \qquad i = 1, ..., n-1 \\ x_0 = 1 & \end{cases}$$

$$\begin{cases} x_{i+1} = \left(1 - 2i^2 \left(\frac{T}{n}\right)^3\right) x_i & i = 1, ..., n-1 \\ x_0 = 1 \end{cases}$$

d) This problem also can be solved explicitly

$$\begin{aligned} & x' + 2t^2 x = 0 \\ \Rightarrow & \frac{x'}{x} = -2t^2 \\ \Rightarrow & (\ln x)' = -2t^2 \\ \Rightarrow & \ln x = -\frac{2}{3}t^3 + \ln x_0 = -\frac{2}{3}t^3 \\ \Rightarrow & x = e^{-\frac{2}{3}t^3}. \end{aligned}$$

6.4.3. Solution to exercise84. a) From Ito's formula we have:

$$dS_t = df(X_t) = \left(\frac{\partial f}{\partial t}(t, X_t) + \frac{\partial f}{\partial x}(t, X_t)\mu(t, X_t) + \frac{1}{2}\sigma^2(t, X_t)\frac{\partial^2 f}{\partial x^2}(t, X_t)\right) dt$$
$$+ \sigma(t, X_t)\frac{\partial f}{\partial x}(t, X_t) dW_t$$

Since $S_t = f(X_t)$, it is easy to find: $\dfrac{\partial f}{\partial t} = 0$, $\dfrac{\partial f}{\partial x} = y(x)$ and $\dfrac{\partial^2 f}{\partial x^2} = \dfrac{dy}{dx} = y'(x)$. Thus, given

$$dX_t = X_t^2 dt + dW_t,$$

we have:

$$dS_t = \left(y(X_t)X_t^2 + \frac{1}{2}y'(X_t)\right) dt + y(X_t) dW_t.$$

Note $y(X_t)X_t^2 + \frac{1}{2}y'(X_t) = 0$, so we have $dS_t = y(X_t)dW_t$. From exercise 83 we knew that $y(x) = e^{-\frac{2}{3}x^3}$. So, $dS_t = e^{-\frac{2}{3}X_t^3} dW_t$.
b) S_t can be discretized as $S_{t+h} = S_t + e^{-\frac{2}{3}X_t^3}\Delta W_t$ where $X_{t+h} = X_t +$

$X_t^2 h + \Delta W_t$. So for $i = 0, 1, 2, ..., n$ one can have:

$$\begin{cases} S_{(i+1)\frac{T}{n}} = S_{i\frac{T}{n}} + e^{-\frac{2}{3}X^3_{i\frac{T}{n}}}\Delta W_{i\frac{T}{n}} \\ X_{(i+1)\frac{T}{n}} = X_{i\frac{T}{n}} + X^2_{i\frac{T}{n}}\frac{T}{n} + \Delta W_{i\frac{T}{n}} \\ X_0 = 0 \end{cases}.$$

c) Using Example 81, or just by repeating the same process, we have $E(S_T) = 0$ and $Var(S_T) = \int_0^T E\left(e^{-\frac{4}{3}X_t^3}\right) dt$.

6.4.4. Solution to exercise 85. a) For using the Ito lemma we need to introduce $f(t,x) = \exp(x)$. Then, $\frac{\partial f}{\partial t} = 0$, $\frac{\partial f}{\partial x} = \exp(x)$ and $\frac{\partial^2 f}{\partial x^2} = \exp(x)$. If we get $y_t = f(t, r_t) = \exp(r_t)$, using the Ito lemma we get

$$\begin{aligned} dy_t &= \left(\frac{\partial f}{\partial t}(t,r_t) + \mu(t,r_t)\frac{\partial f}{\partial x}(t,r_t) + \frac{1}{2}\sigma^2(t,r_t)\frac{\partial^2 f}{\partial x^2}(t,r_t)\right) dt \\ &+ \sigma(t,r_t)\frac{\partial f}{\partial x}(t,r_t)dW_t \\ &= \left(\sin(t)\exp(r_t) + \frac{1}{2}\sigma^2\exp(r_t)\right) dt + \sigma\exp(r_t)\, dW_t \\ &= y_t\left(\sin(t) + \frac{1}{2}\sigma^2\right) dt + \sigma y_t dW_t. \end{aligned}$$

b) Consider a grid $0 = t_0 < t_1 < ... < t_N = T$ where $t_{i+1} - t_i = h = \frac{T}{n}$. Now we approximate dr_t by $\Delta r_t = r_{t+h} - r_t$, dt by h and dW_t by $\Delta W_t = W_{t+h} - W_t$. So we have

$$dr_t \approx \Delta r_t = \sin(t)h + \sigma\Delta W_t.$$

c) We have $r_T = \sum_{i=0}^{N-1} \Delta r_{\frac{iT}{N}} = \sum_{i=0}^{N-1} \left(\sin\left(\frac{iT}{N}\right) h + \sigma \Delta W_{\frac{iT}{N}}\right)$. Observe,

$$
\begin{aligned}
E\left(\exp\left(r_T\right)\right) &= E\left(\exp\left(\sum_{i=0}^{N-1}\left(\sin\left(\frac{iT}{N}\right) h + \sigma \Delta W_{\frac{iT}{N}}\right)\right)\right) \\
&= \exp\left(\sum_{i=0}^{N-1} \sin\left(\frac{iT}{N}\right) h\right) E\left(\exp\left(\sum_{i=0}^{N-1} \sigma \Delta W_{\frac{iT}{N}}\right)\right) \\
&= \exp\left(\sum_{i=0}^{N-1} \sin\left(\frac{iT}{N}\right) h\right) E\left(\exp\left(\sigma W_T\right)\right) \\
&= \exp\left(\sum_{i=0}^{N-1} \sin\left(\frac{iT}{N}\right) h\right) \exp\left(\frac{\sigma^2 T}{2}\right) \\
&\underset{h\to 0}{\longrightarrow} \left(\exp \int_0^T \sin(x)\, dx\right) \exp\left(\frac{\sigma^2 T}{2}\right) \\
&= \exp\left(1 - \cos(T)\right) \exp\left(\frac{\sigma^2 T}{2}\right) \\
&= \exp\left(1 - \cos(T) + \frac{\sigma^2 T}{2}\right).
\end{aligned}
$$

d) We use the discretization of the following process to find the same quantity $dy_t = y_t\left(\sin(t) + \frac{1}{2}\sigma\right) dt + \sigma y_t dW_t$:

$$
\begin{aligned}
E\left(y_{t+h}\right) - E\left(y_t\right) = E\left(\Delta y_t\right) &= E\left(y_t\left(\sin(t) + \frac{1}{2}\sigma^2\right) h + \sigma y_t \Delta W_t\right) \\
&= E\left(y_t\right)\left(\sin(t) + \frac{1}{2}\sigma^2\right) h + \sigma E\left(y_t \Delta W_t\right) \\
&= E\left(y_t\right)\left(\sin(t) + \frac{1}{2}\sigma^2\right) h + \sigma E\left(y_t\right) \underbrace{E\left(\Delta W_t\right)}_{=0} \\
&= E\left(y_t\right)\left(\sin(t) + \frac{1}{2}\sigma^2\right) h.
\end{aligned}
$$

Therefore, we have

$$\frac{E(y_{t+h})-E(y_t)}{h}=E(y_t)\left(\sin(t)+\frac{1}{2}\sigma^2\right).$$

If we denote $z_t=E(y_t)$ then by sending $h\to 0$ we get $z'=z\left(\sin(t)+\frac{1}{2}\sigma^2\right)$. This means

$$\frac{d\log(z)}{dt}=\left(\sin(t)+\frac{1}{2}\sigma^2\right).$$

Taking integral over $[0,T]$ we get $\log(z_T)-\underbrace{\log(z_0)}_{\log(E(y_0))=\log(1)=0}=\int_0^T\left(\sin(t)+\frac{1}{2}\sigma^2\right)dt=$ $1-\cos(T)+\frac{\sigma^2}{2}T$. As a result $z_T=\exp\left(1-\cos(T)+\frac{\sigma^2}{2}T\right)$.

6.4.5. Solution to exercise 86. a) In order to discretize we need to introduce $\Delta X_t=X_{t+h}-X_t$, $\Delta W_t=W_{t+h}-W_t$, and $h=\frac{T}{n}$. So we consider $dX_t\approx\Delta X_t$, $dW_t\approx\Delta W_t$ and $dt\approx h$. The form of the discretized SDE is as follows$\Delta X_t=\mu(t,X_t)h+\sigma(t,X_t)\Delta W_t$. This gives the following on the grid points $X_{\frac{i+1}{n}T}=X_{\frac{i}{n}T}+\mu\left(t,X_{\frac{i}{n}T}\right)h+\sigma\left(t,X_{\frac{i}{n}T}\right)\Delta W_{\frac{i}{n}T}$. We have

$$\begin{aligned}x_{t+h}-x_t&=E(X_{t+h})-E(X_t)\\&=E(\Delta X_t)=E(\mu(t,X_t)h+\sigma(t,X_t)\Delta W_t)\\&=E(\mu(t,X_t))h+E(\sigma(t,X_t)\Delta W_t)\\&=E(\mu(t,X_t))h+E(\sigma(t,X_t))\underbrace{E(\Delta W_t)}_{=0}.\end{aligned}$$

Note that since the increments are independent then ΔW_t and $\sigma(t,X_t)$ are independent and therefore, $E(\sigma(t,X_t)\Delta W_t)=E(\sigma(t,X_t))E(\Delta W_t)$. Also we used the fact that$\Delta W_t\sim N(0,h)$; which gives $E(\Delta W_t)=0$. So, we get $\frac{x_{t+h}-x_t}{h}=E(\mu(t,X_t))$. If we send $N\to\infty$ or $h\to 0$ then we get $x'=E(\mu(t,X_t))$. On the other hand, $x_0=E(X_0)=X_0$. So, the ODE will be $\begin{cases}x'=E(\mu(t,X_t))\\x_0=X_0\end{cases}$.

b) For using the Ito lemma we need to consider $f(t,x)=x^2$. So we have $\frac{\partial f}{\partial t}=0$, $\frac{\partial f}{\partial x}=2x$ and $\frac{\partial^2 f}{\partial x^2}=2$. Using the Ito lemma, we know the dynamic is

$$\begin{aligned} dy_t &= \left(\frac{\partial f}{\partial t}(t,X_t)+\mu(t,X_t)\frac{\partial f}{\partial x}(t,X_t)+\frac{1}{2}\sigma^2(t,X_t)\frac{\partial^2 f}{\partial x^2}(t,X_t)\right)dt \\ &+\sigma(t,X_t)\frac{\partial f}{\partial x}(t,X_t)\,dW_t \\ &= \left(2X_t\mu(t,X_t)+\sigma^2(t,X_t)\right)dt+2X_t\sigma(t,X_t)\,dW_t \end{aligned}$$

Similar to the last parts we have $\Delta y_t=\left(2X_t\mu(t,X_t)+\sigma^2(t,X_t)\right)h+2X_t\sigma(t,X_t)\Delta W_t$, where $\Delta y_t=y_{t+h}-y_t$. The other parts are similar to the previous steps. Like before,

$$\begin{aligned} z_{t+h}-z_t &= E(y_{t+h})-E(y_t) \\ &= E(\Delta y_t)=E\left(\left(2X_t\mu(t,X_t)+\sigma^2(t,X_t)\right)h+2X_t\sigma(t,X_t)\Delta W_t\right) \\ &= E\left(\left(2X_t\mu(t,X_t)+\sigma^2(t,X_t)\right)h\right). \end{aligned}$$

Note, since the increments of the Brownian motion are independent then $2X_t\sigma(t,X_t)$ and ΔW_t are independent and we get $E(2X_t\sigma(t,X_t)\Delta W_t)=E(2X_t\sigma(t,X_t))E(\Delta W_t)=E(2X_t\sigma(t,X_t))\times 0=0$. So, we get finally $\frac{z_{t+h}-z_t}{h}=E\left(2X_t\mu(t,X_t)+\sigma^2(t,X_t)\right)$, and if we send $h\to 0$ then we get $z_t'=E\left(2X_t\mu(t,X_t)+\sigma^2(t,X_t)\right)$. Furthermore, $z_0=E\left(X_0^2\right)=X_0^2$, as a result the ODE is $\begin{cases} z'=E\left(2X_t\mu(t,X_t)+\sigma^2(t,X_t)\right) \\ z_0=X_0^2 \end{cases}$.

c) We have the following cases.

(1) For, $dX_t=-X_tdt+dW_t$, we have $x'=E(-X_t)=-x$ with $x_0=X_0$. Furthermore, we have $z'=E\left(2X_t\mu(t,X_t)+\sigma^2(t,X_t)\right)=E\left(-2X_t^2+1\right)=-2z+1$ with $z_0=X_0^2$.

(2) For $dX_t = (1 - X_t)dt + dW_t$, we have $x' = E(1 - X_t) = 1 - x$ with $x_0 = X_0$. First let us solve this ODE. Indeed, we have

$$\begin{aligned} x' + x = 1 &\Rightarrow e^t (x' + x) = e^t \\ &\Rightarrow \left(xe^t\right)' = e^t \Rightarrow xe^t - X_0 = e^t - 1 \\ &\Rightarrow x = 1 + e^{-t}(X_0 - 1). \end{aligned}$$

Now, for the second ODE we have and

$$\begin{aligned} z' &= E\left(2X_t \mu(t, X_t) + \sigma^2(t, X_t)\right) \\ &= E(2X_t(1 - X_t) + 1) = 2E(X_t) - 2E\left(X_t^2\right) + 1 \\ &= 2x - 2z + 1 = 2\left(1 + e^{-t}(X_0 - 1)\right) - 2z + 1. \end{aligned}$$

So we have $z' + 2z = 3 + 2e^{-t}(X_0 - 1)$.

(3) For $dX_t = (1 - X_t)\,dt + \sqrt{X_t}dW_t$, like the previous one we have $x' = 1 - x$. On the other hand, we have

$$\begin{aligned} z' &= E\left(2X_t \mu(t, X_t) + \sigma^2(t, X_t)\right) \\ &= E(2X_t(1 - X_t) + X_t) = 3E(X_t) - 2E\left(X_t^2\right) \\ &= 3x - 2z = 3\left(1 + e^{-t}(X_0 - 1)\right) - 2z. \end{aligned}$$

So we have $z' + 2z = 3(1 + e^{-t}(X_0 - 1))$.

6.4.6. Solution to exercise 87. 1) Let us consider $f(t, x) = \frac{1}{x}$. For Ito lemma we need to see $\frac{\partial f}{\partial t} = 0$, $\frac{\partial f}{\partial x} = -\frac{1}{x^2}$ and $\frac{\partial^2 f}{\partial x^2} = \frac{2}{x^3}$.

$$\begin{aligned} dS_t &= \left(\frac{\partial f}{\partial t}(t, X_t) + \mu(t, X_t)\frac{\partial f}{\partial x}(t, X_t) + \frac{1}{2}\sigma^2(t, X_t)\frac{\partial^2 f}{\partial x^2}(t, X_t)\right) dt \\ &\quad + \sigma(t, X_t)\frac{\partial f}{\partial x}(t, X_t)\,dW_t \\ &= \left(-(1 - X_t)\frac{1}{X_t^2} + \frac{X_t}{X_t^3}\right) dt - \sqrt{X_t}\frac{1}{X_t^2}dW_t \\ &= \frac{1}{X_t}dt - \frac{1}{X_t^{\frac{3}{2}}}dW_t = S_t dt - (S_t)^{\frac{3}{2}}\,dW_t. \end{aligned}$$

2) Like other problems we have $\Delta y_t = S_{t+h} - S_t = S_t h - (S_t)^{\frac{3}{2}}\,\Delta W_t$. If

we set $z_t = E(S_t)$. As a result

$$\begin{aligned} z_{t+h} - z_t &= E(\Delta S_t) \\ &= E\left(S_t h - (S_t)^{\frac{3}{2}} d\Delta W_t\right) = E(S_t h) - E\left((S_t)^{\frac{3}{2}} \Delta W_t\right) \\ &= h z_t - E\left((S_t)^{\frac{3}{2}}\right) \underbrace{E(\Delta W_t)}_{=0} = h z_t. \end{aligned}$$

Note that we used the fact that increments are independent for Brownian motion and that $\Delta W_t \sim N(0, h)$. By sending $h \to 0$, we get $z' = z$. On the other hand, $z_0 = E(S_0) = \frac{1}{x_0}$. So the ODE is

$$\begin{cases} z' = z \\ z_0 = \frac{1}{x_0} \end{cases}.$$