Financial Engineering

WIRTSCHAFTS
UNIVERSITÄT
WIEN VIENNA
UNIVERSITY OF
ECONOMICS
AND BUSINESS

Zehra Eksi

Institute for Statistics and Mathematics
Dept. of Finance, Accounting and Statistics

Spring 2019

Part I

Introduction

Outline

- Organization
- Financial Engineering
- Literature

Course Info

Title:	Financial Engineering (5195)
Instructor:	*Assoc. Prof. Zehra Eksi*
Contact details:	zehra.eksi@wu.ac.at
Office hours:	Tue $13:30 - 15:00$

- Lectures: Tuesdays from 9:00 to 13:00
- Attendance: mandatory (attend at least %80 of all lectures, i.e., at most one out of seven sessions can be missed)

Assessment and Date of Exams

- Weekly homework assignments (%30):
  - Submission as a group of at most four people (%20)
  - Presentation of solutions in the class by one of the group members (%10)
- Two written exams (%60):
  - mid-term (%20): 26.03.2019, TC.4.01
  - final (%40): 30.04.2019, 9:00-11:00, TC.5.15

Prerequisites and Objective

- Prerequisites:
  - Knowledge in (continuous-time) finance
  - Some knowledge in statistics, probability and stochastic processes
  - Knowledge of a programming language
- Main goal: to become familiar with the essential techniques and tools for *financial engineering*
- Material: lecture slides will be updated continously (available at Learn@WU)

Outline

- Organization
- Financial Engineering
- Literature

What is financial engineering?

Financial engineering is...

an interdisciplinary area consisting of finance, engineering, tools of mathematics and the practice of programming.

The main applications of financial engineering are to:

- portfolio management
- risk management
- financial regulation $\Uparrow$
- structured products $\Downarrow$
- derivatives pricing $\Downarrow$
- trading and execution $\Uparrow$

For $\Downarrow$ visit *http://blogs.reuters.com/emanuelderman/2011/07/07/financial-engineering-as-a-career-part-1/*

Topics to be covered

- Principles of derivatives pricing;
- Principles of Monte Carlo;
- Generating random variables and stochastic processes;
- Simple variance reduction techniques;
- Pricing exotic (Bermudan) options by means of Monte Carlo simulation;
- Construction of yield-curve;
- Applications in risk management.

Outline

- Organization
- Financial Engineering
- Literature

References

- Paul Glasserman [PG] : Monte Carlo Methods in Financial Engineering (2004)
- Rüdiger Seydel [RS]: Tools for Computational Finance (2009)
- Paolo Brandimarte [PM]: Handbook in Monte Carlo Simulation: Applications in Financial Engineering, Risk Management, and Economics (2014)
- Damiano Brigo and Fabio Mercurio [BM]: Interest rate models-theory and practice: with smile, inflation and credit (2007)
- Damir Filipovic [DF]: Term Structure Models (2009)

Part II

Principles of Derivatives Pricing

Outline

- Main ideas
- Approaches to Derivatives Pricing

Derivative Instruments

Definition

A derivative is an instrument whose value is derived from the value of one or more underlying assets.

Some examples:

options (European, American, Bermudan option...); futures; forwards; swaps...

Undelying assets include

stocks; bonds; commodities; currencies; weather; inflation; credit risk...

Pricing of Derivatives

- Pricing derivatives constitute an important place in financial engineering.
- Given the structure of the contract and the price of the underlying, the objective is to find the *fair price*.
- Mostly, the idea of "no arbitrage" yields the *fair price*: The price of a derivative security should be equal to the cost of perfectly replicating the security through trading in other assets.

Three main principles to keep in mind

- P1 If a derivative security can be perfectly replicated through trading in other assets (existence of a self-financing replicating strategy), then the price of the derivative is the cost of replication.
- P2 Discounted asset prices are martingales under a probability measure associated with the choice of discount factor (or numeraire).
- P3 In a *complete* market, any payoff can be replicated through a trading strategy, and the martingale measure associated with a numeraire is unique.

Outline

- Main ideas

- Approaches to Derivatives Pricing

Approaches to derivative pricing

WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

PDE Approach

[P1] together with the given dynamics of the underlying asset lead to a partial differential equation (PDE) that the price of the derivative satisfies.

Risk-Neutral (Martingale) Approach

[P2] gives us a way to express the price of the derivative as the expected present value of the terminal payoff discounted at the risk-free rate.

Naturally, results of the two approaches should coincide !

A more technical link through the two approaches I

Feynman-Kac formula

Consider function $\mu(x)$, $\sigma(x)$, $r(x)$ and some function ϕ on $\mathbb{R}$. Suppose that V(t,x) solves the terminal value problem

$$\frac{\partial V}{\partial t}(t,x) + \mu(x)\frac{\partial V}{\partial x}(t,x) + \frac{1}{2}\sigma(x)^2\frac{\partial^2 V}{\partial x^2}(t,x) = r(x)V(t,x),$$
$$V(T,x) = \phi(x). \tag{1}$$

Then, it holds for $t_0 \leq T$ that

$$V(t_0,x) = \mathbb{E}_x\left(\exp\left(-\int_0^{T-t_0} r(X_s)ds\right)\phi(X_{T-t_0})\right), \tag{2}$$

where X solves the SDE

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t,\ X_{t_0} = x. \tag{3}$$

A more technical link through the two approaches II

Feynman-Kac formula can be used in two ways:

- Compute the expectation (2) in order to solve numerically the PDE in (1).
- We can solve (numerically) the PDE in (1) to compute the expectation in (2) .

Possible problems

- PDE Approach:
  - A solution may not exist when underlying asset price dynamics are complex.
  - Numerical solution may be impractical when number of underlyings for replication is large.
- Risk-neutral Approach:
  - Most of the time it is not possible to calculate the expectation (integral) explicitly.
  - Standard numerical solution techniques may be impractical when number of underlyings is large.

Possible solution...

We can use *Monte Carlo* simulation to compute the expectation numerically.

Part III

Principles of Monte Carlo Simulation

Outline

- Monte Carlo Integration
- Generating Random Variables
- Simulating Poisson Process
- Simulating Brownian Motion
- European Option Pricing
- Variance Reduction

Monte Carlo Integration I

- Suppose we want to compute
$$\Theta = \int_0^1 g(x)dx$$
- It may not be possible to compute analytically.
- We make the observation that
$$\Theta = \mathbb{E}(g(U))$$
where $U \sim U(0,1)$.
- Given a $U(0,1)$ random number generator, this gives a way to estimate Θ via:
  1. Generate IID sample $U_1, U_2, \ldots, U_n$ from $U(0,1)$,
  2. Compute
$$\hat{\Theta}_n = \frac{g(U_1) + g(U_2) + \cdots + g(U_n)}{n}.$$
- Is $\hat{\Theta}_n$ a good estimator of Θ?

Monte Carlo Integration II

Properties of $\hat{\Theta}_n$

$\hat{\Theta}_n$ is an *unbiased* and *consistent* estimator of Θ, i.e.,

- $\mathbb{E}(\hat{\Theta}_n) = \Theta$,
- $\hat{\Theta}_n \to \Theta$ as $n \to \infty$, a.s. This is a direct consequence of the Strong Law of Large Numbers (SLLN).

Recall: SLLN

Let $X_1, X_2, \ldots$ be iid random variables with $\mathbb{E}(X_i) = \mu$ and $var(X_i) = \sigma^2 < \infty$, and define $\bar{X}_n = (1/n)\sum_{i=1}^{n} X_i$. Then, for every $\epsilon > 0$,

$$P(\lim_{n\to\infty} |\bar{X}_n - \mu| < \epsilon) = 1;$$

that is $\bar{X}_n \xrightarrow{a.s.} \mu$

Monte Carlo Integration III

Example

- Compute the integral $\Theta = \int_2^4 (x^3 + x)dx$ by Monte Carlo method with $n = 10000$.
  - Notice that $\Theta = 2\int_2^4 (x^3 + x)\frac{1}{2}dx$.
  - That is, for $X \sim U(2,4)$, we have $\Theta = 2\mathbb{E}(X^3 + X)$
  - Hence we can estimate Θ by generating 10000 IID $U(0,1)$, transforming this into (HOW?) IID $U(2,4)$ random variables $X_1, X_2, \ldots, X_{10000}$ and then computing $\hat{\Theta}_n = \frac{2}{n}\sum_{i=1}^{n}(X_i^3 + X_i)$
- One can actually compute the integral analytically. We have $\Theta = 66$.
- How close are the two results?

Monte Carlo Integration IV

Algorithm: Monte Carlo Integration

Given inputs g, interval (a, b), sample size n

1. for $i = 1 : n$
2. generate $U_i \sim U(0, 1)$
3. transform $X_i \leftarrow (b - a)U_i + a$
4. $Y_i \leftarrow g(X_i)$
5. end for
6. $\hat{\Theta} \leftarrow \frac{(b-a)}{n} \sum_{i=1}^{n} Y_i$

Monte Carlo Error I

Monte Carlo error

The Monte Carlo error (MCE) for a given number of simulation trials n is defined as the difference between the estimate Θ_n and Θ:

$$MCE := \Theta_n - \Theta$$

- The error depends on the sample, hence it is random as well.
- But we can characterize the distribution of the MCE by the help of CLT.

Monte Carlo Error II

Theorem (Central Limit Theorem)

Let $X_1, X_2, \ldots$ *be a sequence of iid random variables with* $\mathbb{E}(X_i) = \mu$ *and* $0 < var(X_i) = \sigma^2 < \infty$. *Define* $\bar{X}_n = (1/n)\sum_{i=1}^{n} X_i$. *Let* $G_n(x)$ *denote the cdf of* $\sqrt{n}(\bar{X}_n - \mu)/\sigma$. *Then, for any* x, $-\infty < x < \infty$,

$$\lim_{n\to\infty} G_n(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy;$$

That is, $\frac{(\bar{X}_n - \mu)}{\sigma/\sqrt{n}}$ *(standardized sample means) has a limiting standard normal distribution.*

Monte Carlo Error III

- From the central limit theorem it follows that the MCE converges in distribution to $N(0, \frac{\sigma}{\sqrt{n}})$.
- The term $\sigma/\sqrt{n}$ referred to as the standard error.
- Notice that cutting the error in half requires to quadruple the number of simulations (n) .
- Adding one decimal place of precision requires 100 times as many simulations.

Multi-Dimensional Monte Carlo Integration

- We consider the problem
$$\Theta = \int_0^1 \int_0^1 g(x,y)dxdy.$$
- Recall that for $U^{(1)}$, $U^{(2)}$ independent $U(0,1)$ random variables we have $f(u^1,u^2) = f_1(u^1)f_2(u^2) = 1$ on $(0,1)^2$.
- Hence, we can write $\Theta = \mathbb{E}(g(U^{(1)},U^{(2)}))$.
- To estimate Θ:
  - generate n of U(0,1) random vectors $(U_i^{(1)},U_i^{(2)})$
  - $\hat{\Theta}_n = \frac{g(U_1^{(1)},U_1^{(2)})+g(U_2^{(1)},U_2^{(2)})+\cdots+g(U_n^{(1)},U_n^{(2)})}{n}$
- $\hat{\Theta}_n$ still preserves the desirable properties.

Monte Carlo integration for more general problems

- Suppose now we want to compute

$$\Theta = \int\int_D g(x,y)f(x,y)dxdy$$

  where $f(x,y)$ is some density on D.
- Hence we have $\Theta = \mathbb{E}(g(X,Y))$ where (X,Y) has the joint density $f(x,y)$.
- To estimate Θ we can generate n random vectors (X,Y) from the joint density $f(x,y)$ and compute

$$\hat{\Theta}_n = \frac{g(X_1,Y_1)+g(X_2,Y_2)+\cdots+g(X_n,Y_n)}{n}$$

Outline

- Monte Carlo Integration
- Generating Random Variables
- Simulating Poisson Process
- Simulating Brownian Motion
- European Option Pricing
- Variance Reduction

Methods for Generating Random Variables I

There are three main methods to generate random variables:

- the inverse transform method
- the composition method
- the acceptance-rejection method

For the Inverse transform method, we mainly make use of the following well-known result:

Theorem (Probability integral transformation)

Let X have cdf $F(x)$ and define the RV Y as $Y = F(X)$. Then Y is uniformly distributed on $(0, 1)$, that is, $P(Y \leq y) = y$, $0 < y < 1$.

Methods for Generating Random Variables II

Inverse Transform Method

- We want to sample from a CDF F, i.e., to generate a random variable X with $\mathbb{P}(X \leq x) = F(x)$
- This method sets $X = F^{-1}(U)$, $U \sim U(0,1)$.
- Hence

$$\begin{aligned}\mathbb{P}(X \leq x) &= \mathbb{P}(F^{-1}(U) \leq x) \\ &= \mathbb{P}(U \leq F(x)) \\ &= F(x).\end{aligned}$$

- If the inverse of F is not well-defined we may set

$$F^{-1}(u) = \inf\{x : F(x) \geq u\}.$$

Methods for Generating Random Variables III

Example: exponential distribution-inverse transform method

- We wish to generate $X \sim exp(\lambda)$.
- We have the cdf $F(x) = 1 - e^{-\lambda x}$, $x \geq 0$
- Hence, $F^{-1}(u) = -\log(1-u)/\lambda$.
- To sample from $exp(\lambda)$:
  - i Generate $U \sim U(0,1)$;
  - ii Set $X = -\log(u)/\lambda$ (WHY?).

Methods for Generating Random Variables IV

Example: discrete distributions-inverse transform method

- Suppose we have a discrete random variable with possible values $c_1 < \cdots < c_n$.
- Let p_i be the probability associated to c_i
- Set $q_0 = 0$, and $q_i = \sum_{j=1}^{i} p_j$, $i = 1, 2, \ldots, n$ (Hence $q_i = F(c_i)$).
- To sample from this distribution
  i generate $U \sim U(0, 1)$
  ii find $K \in \{1, \ldots, n\}$ s.t. $q_{K-1} < U \leq q_K$
  iii set $X = c_K$.

Methods for Generating Random Variables V

The Composition Method

- Suppose we have $X \sim F$ and we can write $F(x) = \sum_{i=1}^{\infty} w_i F_i(x)$, where $w_i \geq 0$ and $\sum_i w_i = 1$ and F_is are cdfs.
- We may often have such representations, e.g., $\mathit{Hyperexp}(\lambda_1, \alpha_1 \ldots, \lambda_n, \alpha_n)$ with

$$f(x) = \sum_{i=1}^{n} \alpha_i \lambda_i e^{-\lambda_i x}$$

- *How can we show that this method actually works?*
- We can make use of the following algorithm:
  - i Generate K that is distributed on the positive integers s.t $\mathbb{P}(K = j) = w_j$. (How can we do this?)
  - ii If $K = j$, then generate Z_j from the cdf F_j;
  - iii Set $X = Z_j$.

Methods for Generating Random Variables VI

Acceptance-Rejection Method

- Suppose we want to generate sample for a rv X with density f and cdf, F.
- Suppose it's hard to simulate a value of X directly using inverse transform or composition algorithms.
- Let Y be another rv with density g and suppose it's easy to simulate Y.
- If there exists a constant c such that $f(x) \leq cg(x)$, for all x, then we can simulate a value of X as:

  i generate Y from distribution g

  ii generate $U \sim U(0,1)$

  iii if $U \leq f(Y)/cg(Y)$

  return X

  otherwise

  go to Step(i).

Methods for Generating Random Variables VII

Generating Multivariate Normals

- Suppose we want to generate the random vector $X = (X_1, \ldots, X_n)$ where $X \sim N_n(0, \Sigma)$.
- Let $Y = (Y_1, \ldots, Y_n)$ where Y_is are IID $N(0, 1)$.
- If A is an $n \times n$ matrix then

$$Z = AY \sim N_n(0, AA^\top)$$

- We can generate independent Normal rvs $Y_1, \ldots, Y_n$ and consider them as a vector.
- Thus, the problem of sampling from X reduces to finding a matrix A s.t. $AA^\top = \Sigma$

Methods for Generating Random Variables VIII

Cholesky factorization

- Among all possible A, a lower triangular one is obtained as a result of Cholesky factorization
- However, be careful if Σ is positive semi-definite (hence rank deficient).
- In this case it is better to reduce the problem to one of full rank, find subvector $\tilde{X}$ and matrix D s.t the covariance matrix $\tilde{\Sigma}$ is full rank and that

$$D\tilde{X} = \Sigma.$$

- Cholesky factorization can now be applied to $\tilde{\Sigma} = \tilde{A}\tilde{A}^\top \Rightarrow X = D\tilde{A}Y$.
- Such a situation may occur. e.g., in case of factor models in which the vector X of length n is determined by $k < d$ number of risk sources (factors).

Outline

- Monte Carlo Integration
- Generating Random Variables
- Simulating Poisson Process
- Simulating Brownian Motion
- European Option Pricing
- Variance Reduction

Simulating Poisson Process I

Poisson Process

Let $(\tau_i)_{i\geq 1}$ be a sequence of independent exponential random variables with parameter λ and $T_n = \sum_{i=1}^{n} \tau_i$. The process $\{N_t, t \geq 0\}$ defined by

$$N_t = \sum_{n\geq 1} I_{\{t\geq T_n\}}$$

is called a Poisson process with intensity λ.

Simulating Poisson Process II

- For a Poisson process the numbers of arrivals in non-overlapping intervals are independent and the distribution of the number of arrivals in an interval only depends on the length of the interval.
- It is a counting process with

$$\mathbb{P}(N_t = r) = \frac{(\lambda t)^r e^{-\lambda t}}{r!}.$$

- From its definition, one can simulate a Poisson process by simply generating the $exp(\lambda)$ inter-arrival times, τ_i.

Simulating Poisson Process III

Simulation Algorithm: Poisson Process

set $t = 0$, $I = 0$

generate $U \sim U(0, 1)$

set $t = t - \log(U)/\lambda$

while $t < T$

    set $I = I + 1$, $S(I) = t$
    generate $U \sim U(0, 1)$
    set $t = t - \log(U)/\lambda$

Outline

- Monte Carlo Integration
- Generating Random Variables
- Simulating Poisson Process
- Simulating Brownian Motion
- European Option Pricing
- Variance Reduction

One-dimensional standard Brownian motion

Brownian motion

One-dimensional standard Brownian motion on $[0, T]$ is a stochastic process $\{W(t), 0 \leq t \leq T\}$ with the following properties:

i. $W(0) = 0$;

ii. the mapping $t \mapsto W(t)$ is, with probability 1, a continuous function on $[0, T]$;

iii. the increments $\{W(t_1) - W(t_0), W(t_2) - W(t_1), \cdots, W(t_k) - W(t_{k-1})\}$ are independent for any k and any $0 \leq t_0 < t_1 < \cdots < t_k \leq T\}$,

iv. $W(t) - W(s)$ is distributed as $N(0, t-s)$ for any $0 \leq s < t \leq T$.

Note that from i. and iv. $W(t) \sim N(0, t)$. Also for constants μ and $\sigma > 0$, we call process (X_t) a Brownian motion with drift μ and diffusion coefficient σ^2 if

$$\frac{X(t) - \mu t}{\sigma}$$

is a standard BM by setting $X(t) = \mu t + \sigma W(t)$.

Random walk construction

- In discussing the simulation of BM, we can focus on simulating values $(W_{t_0}), \ldots, W_{t_n})$ at a fixed set of points $0 < t_1 < \ldots < t_n$, since BM has independent normally distributed increments.
- Let $Z_1, \ldots, Z_n$ be independent standard normal variables, then for a standard BM, with $t_0 = 0$ and $W(0) = 0$, we can generate

$$W(t_{i+1}) = W(t_i) + \sqrt{t_{i+1} - t_i} Z_{i+1}, \quad i = 0, \ldots, n-1.$$

- The vector $(W_{t_1}), \ldots, W_{t_n})$ is a linear transformation of the vector of increments $\{W(t_1) - W(t_0), W(t_2) - W(t_1), \cdots, W(t_n) - W(t_{n-1})\}$, and since these increments are independent and normally distributed we can conclude that $(W_{t_1}), \ldots, W_{t_n})$ has a multivariate normal distribution.

Simulation with Cholesky Factorization

- Note that for simulating the multivariate normal, we need mean vector and the covariance matrix.
- From the independent increments property one can show that for $s \leq t$ $Cov(W(s), W(t)) = s$, and let C denote the covariance matrix of $(W_{t_1}, \ldots, W_{t_n})$, with the entries $C_{ij} = \min(t_i, t_j)$.
- $(W_{t_1}, \ldots, W_{t_n})$ has the distribution $N(0, C)$ and one can simulate this vector as AZ, where $Z = (Z_1, \ldots, Z_n)^\top \sim N(0, I)$ and A satisfies $AA^\top = C$
- The Cholesky factorization for C yields the lower triangular matrix A given by

$$A = \begin{bmatrix} \sqrt{t_1} & 0 & \cdots & 0 \\ \sqrt{t_1} & \sqrt{t_2 - t_1} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \sqrt{t_1} & \sqrt{t_2 - t_1} & \cdots & \sqrt{t_n - t_n - 1} \end{bmatrix},$$

Simulation of Geometric Brownian Motion

Definition:Geometric Brownian Motion

A stochastic process $\{X_t : t \geq 0\}$, is a geometric Brownian motion (GBM) with drift μ and volatility σ if

$$\log(X) \sim BM(\mu - \frac{\sigma^2}{2}, \sigma).$$

That is

$$X_t \sim \log N((\mu - \frac{\sigma^2}{2})t, \sigma^2 t)$$

- Question: How would you simulate X_{t_i}?

Impact of volatility

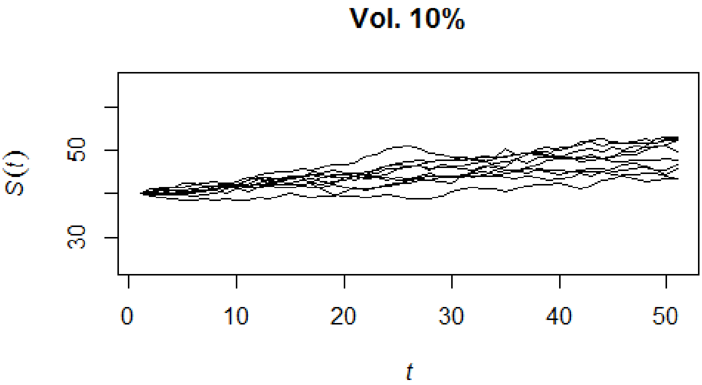

Figure: Generated paths for a GBM with $S_0 = 40\ \mu = 0.25,\ \sigma = 0.1$

Impact of volatility

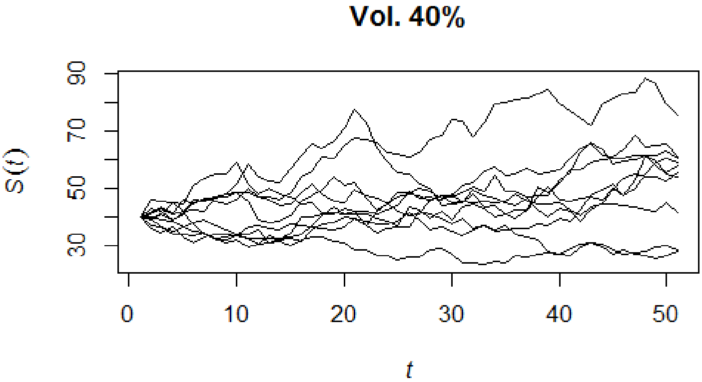

Figure: Generated paths for a GBM with $S_0 = 40$ $\mu = 0.25$, $\sigma = 0.4$ (We fix the seed to see the impact of volatility)

Outline

- Monte Carlo Integration
- Generating Random Variables
- Simulating Poisson Process
- Simulating Brownian Motion
- European Option Pricing
- Variance Reduction

Application: Pricing Standard European Options I

Option payoffs

- European call option written on S:

$$V(S_T) = \max(S_T - K, 0) = (S_T - K)^+$$

- European put option written on S:

$$V(S_T) = \max(K - S_T, 0) = (K - S_T)^+$$

Application: Pricing Standard European Options II

The Black-Scholes Model

- Suppose we are given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.
- We assume that we have a frictionless market without any arbitrage opportunities and continuous trading over $[0, T]$.
- There are two main assets traded in the market:
  - Risk-free bond: $B_t = B_0 e^{rt}$, $r \geq 0$ is the constant risk-free rate
  - Risky stock: follows GBM dynamics, that is

$$S_t = S_0 \exp\left((\mu - \sigma^2/2)t + \sigma\sqrt{t}W_t\right),$$

    where μ is the drift, σ is the volatility and W is an $\mathbb{F}$-Brownian motion.
- Our objective is to come up with the $t = 0$ price, C, of a call option written on the stock with strike price K.

Application: Pricing Standard European Options III

WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

- It follows from no-arbitrage arbitrage assumption that

$$C = \mathbb{E}^{\mathbb{Q}}(e^{-rT}(S_T - K)^+). \tag{4}$$

- Notice that the expectation is taken under the so-called *martingale* or *risk-neutral* probability measure $\mathbb{Q}$.
- This implies that in our analytical and numerical calculations we need the *risk-neutral* dynamics of the stock prices.
- By choosing the market price of risk (Girsanov density kernel or Radon-Nikodym derivative) $\lambda = \frac{\mu - r}{\sigma}$ we can change the measure from $\mathbb{P}$ to $\mathbb{Q}$ under which $W_t^{\mathbb{Q}} = W_t + \lambda t$ is a $\mathbb{Q}$-Brownian motion.
- This yields, as desired, that the discounted stock price is a martingale with the dynamics

$$e^{-rt}S_t = S_0 \exp\left(-\frac{\sigma^2}{2}T + \sigma\sqrt{T}W_t^{\mathbb{Q}}\right).$$

Application: Pricing Standard European Options IV

Closed-form Price

Under the Black-Scholes model, price of the call option with strike K is given by

$$C = S_0\Phi(d_1) - e^{-rT}K\Phi(d_2) \tag{5}$$

where

$$d_1 = \frac{\log(S_0/K) + (r + \sigma^2/2)T}{\sigma\sqrt{T}}$$
$$d_2 = d_1 - \sigma\sqrt{T}$$

Application: Pricing Standard European Options V

Numerical Valuation

- As an alternative, we can rely on the numerical computation of the expectation in (4).
- To this, we can use Monte Carlo. That is, the estimator

$$\hat{C}_n = \frac{1}{n}\sum_{i=1}^{n} e^{-rT}(S_T^i - K)^+.$$

- Here we need to simulate S_T^is (under risk neutral measure) and we know how to do this (see, Simulation of GBM part).

Application: Pricing Standard European Options VI

Agorithm:Pricing European call option

Given inputs S_0, r, σ, K, T, n :number of simulations

1. for i=1:n
2. Generate $Z_i \sim Normal(0, 1)$
3. $S_i \leftarrow S_0 \exp\left((r - \sigma^2/2)T + \sigma\sqrt{T}Z_i\right)$
4. $C_i \leftarrow e^{-rT}(S_i - K)^+$
5. end
6. $\hat{C}_n = \frac{1}{n}\sum_{i=1}^{n} C_i$.

Application: Pricing Standard European Options VII

Example

Suppose we want to price a European call option written on a stock with initial value $S_0 = 100$, $\sigma = 0.3$, $\mu = 0.2$. The maturity of the option is in $T = 1$ year and the strike price is $K = 110$. Assume that the risk-free interest rate is $r = 2\%$. Price the option analytically and numerically (simulate $n = 10000$ paths). Compute the corresponding Monte Carlo standard error .

Analytical pricing:

Plugging in the parameters into B-S option pricing formula given in (5), we obtain $C = 8.864156$.

Application: Pricing Standard European Options VIII

Numerical pricing:

- We take $n = 10000$ and use the pricing algorithm. We obtain $\hat{C}_{10000} = 8.6799276$
- How to estimate the standard error?
  1. First we estimate the standard deviation σ:

$$\hat{\sigma}_n = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(C_i - \hat{C}_n)^2},$$

     where C_i is the price, corresponding to the generated path i.
  2. Hence, $\hat{SE}_n = \frac{\hat{\sigma}_n}{\sqrt{n}}$.
  3. Using this methodology we obtain $\hat{SE}_{10000} = 0.1861706$.

Monte Carlo Recipe for Pricing

1. Replace the drifts of the underlying processes with the risk-free interest rate.
2. Simulate paths of the underlying processes.
3. Calculate the payoff of the derivative security on each path.
4. Discount the payoffs at the risk-free rate.
5. Calculate the average over all paths.

Outline

- Monte Carlo Integration
- Generating Random Variables
- Simulating Poisson Process
- Simulating Brownian Motion
- European Option Pricing
- Variance Reduction

MC Error revisited I

- We have $\hat{\Theta}_n = \frac{1}{n}\sum_{i=1}^{n} Y_i$.
- Denote by $\sigma^2 = Var(Y_i)$. CLT implies that $\frac{\hat{\Theta}_n - \Theta}{\sigma/\sqrt{n}} \to N(0,1)$, as $n \to \infty$.
- How can we construct a $100(1-\alpha)\%$ confidence interval for Θ?
  - Let $z_{1-\alpha/2}$ be the $(1-\alpha/2)$ percentile of the $N(0,1)$ distribution.
  - We have

$$\mathbb{P}\left(-z_{1-\alpha/2} \leq \frac{\sqrt{n}(\hat{\Theta}_n - \Theta)}{\sigma} \leq z_{1-\alpha/2}\right) \approx (1-\alpha)$$

$$\mathbb{P}\left(\hat{\Theta}_n - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \Theta \leq \hat{\Theta}_n + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right) \approx (1-\alpha)$$

  - Note that we can estimate σ^2 via $\hat{\sigma}_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \hat{\Theta}_n)^2$.

MC Error revisited II

- We have the *width* of the confidence interval given by $\frac{2\hat{\sigma}_n z_{1-\alpha/2}}{\sqrt{n}}$.
- We would like the width to be small.
- For a fixed α, we have to increase n if we are to decrease the width of the confidence interval.
- In particular width of the confidence interval decreases according to a square-root law involving $\sqrt{n}$, which is rather bad news!
- Increasing the number of replications is less and less effective, and this brute force strategy may result in a remarkable computational burden.
- Also, $Var(Y_i)$ could be too large, or too much computational cost might be required to simulate Y_is (n is necessarily small).
- Hence an alternative strategy to reduce the width is to adopt a clever sampling strategy in order to reduce the variance $Var(Y_i)$.

Variance Reduction Methods

Among others, we will cover mainly the following techniques:

- Antithetic Variates
- Control Variates
- Conditional Monte Carlo
- Importance Sampling

Antithetic Variates I

Idea

- The antithetic sampling does not necessitate deep knowledge about the problem.
- We want to estimate $\Theta = \mathbb{E}(Y)$.
- Suppose we have a 2-sample (Y_1, Y_2), (Y_is) identically distributed (not necessarily independent). Then $\hat{\Theta} = \frac{Y_1+Y_2}{2}$. This yields

$$
\begin{aligned}
Var(\hat{\Theta}) &= \frac{Var(Y_1) + Var(Y_2) + 2Cov(Y_1, Y_2)}{4} \\
&= \frac{Var(Y_1)}{2}(1 + \rho(Y_1, Y_2)).
\end{aligned}
$$

- In crude Monte Carlo, we generate a sample consisting of independent observations. However, inducing some correlation may be helpful in reducing the variance of Θ.

Antithetic Variates II

Uniform Antithetic Variates

- Suppose we have Y a function of IID $U(0,1)$ rvs, s.t., $\Theta = \mathbb{E}(h(U))$ where $U = (U^1, \dots, U^m)$.
- We can construct an estimate for Θ as follows:
  - Set $Y_i = h(U_i)$, where $U_i = (U_i^1, \dots, U_i^m)$.
  - Also set $\overline{Y}_i = h(1 - U_i)$, where $1 - U_i = (1 - U_i^1, \dots, 1 - U_i^m)$.
  - Set $Z_i = \frac{(Y_i + \overline{Y}_i)}{2}$
  - Then, we come up with the estimator

$$\hat{\Theta}_{A,n} = \frac{1}{n} \sum_{i=1}^{n} Z_i.$$

- U^i and $1 - U^i$ are *antithetic variates* and $\hat{\Theta}_{A,n}$ is unbiased and consistent.

Antithetic Variates III

Which algorithm is better? Variance Comparison

Since we have effectively doubled the sample size, we must compare $Var(\hat{\Theta}_{A,n})$ against the variance $Var(\hat{\Theta}_{2n})$ of an independent sample of size $2n$:

$$Var(\hat{\Theta}_{2n}) = Var\left(\frac{\sum_{i=1}^{2n} Y_i}{2n}\right) = \frac{Var(Y_i)}{2n}.$$

$$\begin{aligned} Var(\hat{\Theta}_{A,n}) &= Var\left(\frac{\sum_{i=1}^{n} Z_i}{n}\right) = \frac{Var(Z_i)}{n} \\ &= \frac{Var(Y_i + \overline{Y}_i)}{4n} = \frac{Var(Y_i)}{2n} + \frac{cov(Y_i, \overline{Y}_i))}{2n} \\ &= Var(\hat{\Theta}_{2n}) + \frac{cov(Y_i, \overline{Y}_i))}{2n} \end{aligned}$$

Antithetic Variates IV

Variance Comparison

We have seen

$$Var(\hat{\Theta}_{A,n}) < Var(\hat{\Theta}_{2n}) \Leftrightarrow cov(Y_i, \overline{Y}_i)) < 0.$$

Recall that we have $Y = h(U)$ and $\overline{Y} = h(1 - U)$. Following sufficient condition on h guarantees the desired variance reduction.

Theorem (Variance Comparison, a sufficent condition)

Suppose $h(u^1, ...u^m)$ is a monotonic function of each of its arguments on $[0,1]^m$, then for a set $U = (U^1, \ldots, U^m)$ of IID $U(0,1)$ random variables it holds that

$$Cov(h(U), h(1 - U)) < 0.$$

Antithetic Variates V

Non-Uniform Antithetic Variates

- Consider the case $\Theta = \mathbb{E}(Y)$ where $Y = h(X_1, \ldots, X_m)$, and where $(X_1, \ldots, X_m)$ is a vector of independent random variables.
- If we can make use of the inverse transform method to generate the X_is, we can use antithetic variable method for such problems:
  - Suppose $X_i \sim F_i$
  - Generate $U_1, \ldots, U_m \sim$ IID $U(0, 1)$
  - Set $Z = h(F_1^{-1}(U_1), \ldots, F_m^{-1}(U_m))$
- Since the CDF of any random variable is non-decreasing, it follows that F_i^{-1} also non-decreasing.
- So if, e.g., h is monotonic, so does $h(F^{-1}(\cdot))$ and antithetic variates method works.

Antithetic Variates VI

Normal Antithetic Variates

- Recall that we can not apply inverse transform method to Normal rvs. Still, we can generate antithetic normal random variates.
- Suppose $X \sim N(\mu, \sigma)$. Let $\overline{X} = 2\mu - X$. Then $\overline{X} \sim N(\mu, \sigma)$.
- We have X and $\overline{X}$ negatively correlated. Indeed:

$$\rho(X, \overline{X}) = \frac{Cov(X, \overline{X})}{\sqrt{\sigma^2 \sigma^2}} = \frac{cov(X, -X)}{\sigma^2} = \frac{-\sigma^2}{\sigma^2} = -1$$

- So if $\Theta = \mathbb{E}(h(X_1, \ldots, X_m))$ where the X_i 's $\sim$ IID $N(\mu, \sigma)$ and $h(\cdot)$ is monotonic, then we can again achieve a variance reduction by using antithetic variates.
- How would you use antithetic variate method to price a simple European call option?

Antithetic Variates VII

Example

Use plain Monte Carlo integration and Antithetic sampling to estimate

$$\Theta = \int_0^1 e^x dx$$

- The true value is $e - 1 \approx 1.7183$
- For $\alpha = 0.05$, i.e., 95% confidence level, MC integration yields (MCestm,SEMC,LBMC,UBMC,MCwidth):

$$(1.7170,\, 0.0482,\, 1.6225,\, 1.8116,\, 0.1891)$$

- The estimated value is quite close to the true one. For another seed we could get a much larger or smaller estimate! The width of CI suggests that a small sample consisting of only 100 observations does not yield a reliable estimate.

Antithetic Variates VIII

Example cont...

- For a fair comparison we consider 50 antithetic pairs $(U_i, 1 - U_i)$
- We set $Z_i = \frac{exp(U_i)+exp(1-U_i)}{2}$
- For $\alpha = 0.05$, i.e., 95% confidence level, and fixed seed, AV sampling yields (AVestm,SEAV,LBAV,UBAV,AVwidth):

$$(1.7145,\ 0.0074\ 1.7001\ 1.7289\ 0.02881)$$

- Now the confidence interval is much smaller and, despite the limited sample size, the estimate is fairly reliable.

Control Variates I

- AV is easy to apply and it works under the monotonicity assumption. Better results might be obtained by taking advantage of deeper, domain-specific knowledge.
- Suppose we want to estimate $\Theta = \mathbb{E}(X)$, and that there is another random variable Y, with a known expected value ν, which is correlated with X.
- For example, Θ can be the unknown price of an option, and ν could be the price of a corresponding vanilla option.
- The variable Y is called the *control variate*.
- The correlation with Y may be exploited by adopting the controlled estimator:

$$\hat{\Theta}_c = X + c(Y - \mathbb{E}(Y)) = X + c(Y - \nu),$$

where c is a parameter that we must choose.

Control Variates II

How to choose c?

- We have
  - $\mathbb{E}(\hat{\Theta}_c) = \Theta$, i.e., $\hat{\Theta}_c$ is an unbiased estimator, $\forall c$.
  - $Var(\hat{\Theta}_c) = Var(X) + c^2 Var(Y) + 2c Cov(X, Y)$
- By a suitable choice of c, we could minimize the variance of estimator:

$$c^* = -\frac{Cov(X, Y)}{Var(Y)}$$

- This yields:

$$Var(\hat{\Theta}_{c^*}) = Var(X) - \frac{Cov(X, Y)^2}{Var(Y)} = Var(\hat{\Theta}) - \frac{Cov(X, Y)^2}{Var(Y)}$$

- Hence there is a room for variance reduction when $Cov(X, Y) \neq 0$.

Control Variates III

Estimation of c

- Problem: In practice, the optimal value of c must be estimated, since $Cov(X,Y)$ and possibly $Var(Y)$ are not known.
- Solution: We run k pilot simulations to estimate unknowns:

$$\widehat{Cov}(X,Y) = \frac{\sum_{i=1}^{k}(X_i - \hat{X}_k)(Y_i - \nu)}{k-1}, \quad \widehat{Var}(Y) = \frac{\sum_{i=1}^{k}(Y_i - \nu)^2}{k-1}.$$

- Finally we obtain

$$\hat{c}^* = -\frac{\widehat{Cov}(X,Y)}{\widehat{Var}(Y)}.$$

Control variates method yields the estimator:

$$\hat{\Theta}_{c^*} = \frac{\sum_{i=1}^{n}(X_i + c^*(Y_i - \nu))}{n}$$

Control Variates IV

Algorithm: Control variates

Pilot Simulation

1. for $i = 1 : k$
2. generate (X_i, Y_i)
3. end
4. $\hat{c}^* \leftarrow -\frac{\widehat{Cov}(X,Y)}{\widehat{Var}(Y)}$

Main Simulation

1. for $i = 1 : n$
2. generate (X_i, Y_i), set $Z_i \leftarrow X_i + \hat{c}^*(Y_i - \nu)$
3. end
4. Set $\hat{\Theta}_{\hat{c}^*} \leftarrow \frac{1}{n}\sum_{i=1}^{n} Z_i$.

Note: One should not merge the pilot and main steps! (Bias since c* becomes a random variable depending on X itself).

Control Variates V

Example

We want to estimate $\Theta = \int_0^1 e^x dx = \mathbb{E}(e^U)$, $U \sim U(0,1)$.

- Let us choose the control variate $Y \sim U(0,1)$. Hence $\nu = 0.5$ and $\rho(e^U, Y) = 0.994$
- We set $Z = e^U + c^*(Y - 0.5)$.
- Pilot step with $n = 50$ yields $\hat{c}^* = -1.679454$
- The main step yields $(CVestm, SECV, LBCV, UBCV, CVwidth) =$

$$(1.7132,\ 0.0049,\ 1.7035,\ 1.7228,\ 0.01927).$$

- Compared to naive estimator we observe a remarkable reduction in the variance. This is mostly die to the strong correlation between e^U and Y.
- A more interesting example is to price vanilla call option by taking the stock price value at maturity as a control variate...

Conditional Monte Carlo I

- The idea is simple: we use our knowledge about the problem being studied to reduce the variance of our estimator.
- We want to estimate $\Theta = \mathbb{E}(X)$ where $X = (X_1, \ldots, X_m)$.
- Suppose $Y = h(X)$ and we set $V = \mathbb{E}(Y|Z)$. V is a rv that depends on Z, hence we can write $V = g(Z)$, for some $g(\cdot)$.
- Law of iterated expectations:

$$\mathbb{E}(V) = \mathbb{E}(\mathbb{E}(Y|Z)) = \mathbb{E}(Y)$$

- Hence in order to estimate Θ we may simulate V instead of Y.

Conditional Monte Carlo II

Variance comparison:

- Suppose Z can be simulated easily and $V = \mathbb{E}(Y|Z)$ can be computed exactly.
- Recall the conditional variance formula:

$$Var(Y) = \mathbb{E}(Var(Y|Z)) + Var(\mathbb{E}(Y|Z)).$$

- We have

$$Var(Y) \geq Var(\mathbb{E}(Y|Z)) = Var(V).$$

- Hence V is a better estimator of Θ than Y.
- Note that in order for the conditional expectation method to work, Y and Z should be dependent (Why?).

Conditional Monte Carlo III

Algorithm: Conditional Monte Carlo

1. for $i = 1 : n$
2. generate Z_i
3. compute $g(Z_i) = \mathbb{E}(Y|Z_i)$
4. set $V_i = g(Z_i)$
5. end
6. set $\hat{\Theta}_{CM} = \frac{1}{n}\sum_{i=1}^{n} V_i$

Conditional Monte Carlo IV

Example

- Suppose we want to estimate $\Theta = \mathbb{P}(X + Y > 4)$ where $X \sim exp(1)$ and $Y \sim exp(1/2)$.
- Let $Z = 1_{\{X+Y>4\}}$. Then we can write $\Theta = \mathbb{E}(Z)$.
- We can estimate Θ via naive MC as follows:
  1. Generate $(X_1, \ldots, X_n)$ and $(Y_1, \ldots, Y_n)$.
  2. Set $Z_i = 1_{\{X_i+Y_i>4\}}$, $i = 1, \ldots, n$.
  3. Set $\hat{\Theta}_n = \frac{1}{n}\sum_{i=1}^{n} Z_i$.

Conditional Monte Carlo V

Example, cont...

We can also solve the problem by conditional Monte Carlo:

- We set $V = \mathbb{E}(Z|Y)$. Then,

$$\mathbb{E}(Z|Y = y) = \mathbb{P}(X + Y > 4|Y = y) = \mathbb{P}(X > 4 - y) = 1 - F_X(4 - y).$$

- Since $X \sim exp(1)$, we have $1 - F_X(4 - y) = e^{-(4-y)}$, if $0 \leq y \leq 4$ and 1 for $y > 4$.
- Hence $V = \mathbb{E}(Z|Y) = e^{-(4-Y)}$, if $0 \leq Y \leq 4$ and 1 for $Y > 4$.
- Conditional Monte Carlo algorithm for estimating Θ is:
  1. Generate $(Y_1, \dots, Y_n)$.
  2. Set $V_i = V = \mathbb{E}(Z|Y_i)$, $i = 1, \dots, n$
  3. set $\hat{\Theta}_{CM} = \frac{1}{n} \sum_{i=1}^{n} V_i$

Importance Sampling I

- We want to estimate $\Theta = \mathbb{E}(h(X))$, where $X \sim f$.
- Suppose g is another density with the property that $g > 0$ whenever $f > 0$.
- We have
$$\Theta = \mathbb{E}(h(X)) = \int \frac{h(x)}{g(x)} f(x) g(x) dx = \mathbb{E}_g \left(\frac{h(X) f(X)}{g(X)} \right)$$
- Naive Monte Carlo generates n samples of X from f and yields:
$$\hat{\Theta}_n = \frac{1}{n} \sum_{i=1}^{n} h(X_i).$$
- Alternatively, we can generate n of X values from g and obtain:
$$\hat{\Theta}_n^{IS} = \frac{1}{n} \sum_{i=1}^{n} \frac{h(X_i) f(X_i)}{g(X_i)}.$$

Importance Sampling II

- $\hat{\Theta}_n^{IS}$ is an unbiased estimator:

$$\mathbb{E}_g(\hat{\Theta}_n^{IS}) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}_g\left(\frac{h(X_i)f(X_i)}{g(X_i)}\right) = \mathbb{E}_g\left(\frac{h(X)f(X)}{g(X)}\right) = \mathbb{E}(h(X)) = \Theta.$$

Variance comparison

Denote by $H(x) = \frac{h(x)f(x)}{g(x)}$. We have the variance

$$Var_g(H(X)) = \int H(x)^2 g(x)dx - \Theta^2 = \int \frac{h(x)^2 f(x)}{g(x)} f(x)dx - \Theta^2.$$

On the other hand

$$Var(h(X)) = \int h(x)^2 f(x)dx - \Theta^2$$

Importance Sampling III

- Hence the *reduction* in variance is

$$Var(h(X)) - Var_g(H(X)) = \int h(x)^2 \left(1 - \frac{f(x)}{g(x)}\right) f(x)dx$$

- We want the reduction to be positive
- Let us denote by L the region in he support of f where $h(x)^2 f(x)$ is large.
- For reduction to be positive we would like to choose g so that $f(x)/g(x)$ is small whenever x is in L.
- That is, we would like a density g which puts more weight on L (importance sampling).

Importance Sampling IV

How to choose g:

- Suppose we choose $g(x) = h(x)f(x)/\Theta$. Then $Var_g(H(X)) = 0$, zero variance estimator! This is not feasible in practice since we do not know Θ and therefore don't know g either. Still, this observation can guide us.
- If we could choose g such that it is *similar* to $h(\cdot)f(\cdot)$, then we might reasonably expect to obtain a large variance reduction.
- Similar could mean to choose g so that it has a similar shape to $h(\cdot)f(\cdot)$.
- In particular, we could try to choose g so that $g(x)$ and $h(x)f(x)$ both take on their maximum values at the same value, say x^*.
- Often g is taken to be from the same family of distributions as f.

Part IV

Simulation of SDEs

Outline

- The Euler Scheme
- The Milstein Scheme
- Improvements and Extensions

The Euler Scheme I

Suppose we have an SDE of the form

$$dS_t = a(t, S_t)dt + b(t, S_t)dW_t$$

Suppose, e.g., we want to simulate values of S_T. We may or may not know the distribution. So simulate a discretized version of the SDE

$$\hat{S}_0, \hat{S}_h, \hat{S}_{2h}, \ldots, \hat{S}_{mh},$$

where m is the number of time steps, h is a constant step size and $m = \lfloor T/h \rfloor$. We write the SDE in the integral form:

$$S_t = S_0 + \int_0^t a(u, S_u)du + \int_0^t b(u, S_u)dW_u.$$

The Euler Scheme II

The idea of Euler scheme is to approximate integrals over $(k-1)h$ to kh by freezing the integrand functions to their value at $(k-1)h$. We have

$$\int_{(k-1)h}^{kh} a(u, S_u)du \approx a((k-1)h, S_{(k-1)h})h \tag{6}$$

$$\int_{(k-1)h}^{kh} b(u, S_u)dW_u \approx b((k-1)h, S_{(k-1)h})(W_{kh} - W_{(k-1)h}) \tag{7}$$

Euler approximation:

$$\hat{S}_{kh} = \hat{S}_{(k-1)h} + a\left((k-1)h, \hat{S}_{(k-1)h}\right) + b\left((k-1)h, \hat{S}_{(k-1)h}\right)\sqrt{h}Z_k,$$

where Z_ks are IID $N(0,1)$.

The Euler Scheme III

Even though we only care about S_T, we still need to generate intermediate values, S_{ih}, if we are to minimize the discretization error:

- This means that simulating SDEs is computationally intensive.
- Because of the discretization error, $\hat{\Theta}_n$ is no longer an unbiased estimator of Θ.
- In general, if we have path dependency, i.e., $\Theta = \mathbb{E}(f(S_{t_1}, \ldots, S_{t_K})$ then we would need to keep track of $(S_{t_1}, \ldots, S_{t_K})$.

this is not an unbiased estimator: the bias in this case arises due to discretization

The Euler Scheme IV

Euler scheme: multi-dimensional case

We can generalize this idea into the multidimensional case, $S_t \in \mathbb{R}^d$. Multidimensional case may occur when we have:

- Modeling the evolution of multiple stocks.
- Modeling the evolution of a single stock in a stochastic volatility model.
- Modeling the evolution of interest rates in short rates

If the Brownian motions, W_t , are correlated then we can use the Cholesky decomposition. But most of the time we have standard multi-dimensional Brownian motion (any correlations between components of S_t is presented through induced through $b(t, S_t)$).

The Euler Scheme V

WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

Weak and Strong Order Criterion

Two approaches for measuring the error in a discretization scheme:

- A strong error criterion:

$$\mathbb{E}\left(\left\|\hat{S}_{mh} - S_T\right\|\right)$$

- A weak error criterion:

$$\left|\mathbb{E}\left(f(\hat{S}_{mh}) - f(S_T)\right)\right|,$$

more important in finance as we are interested in payoffs..
but for example for options, the f() is not smooth because of the kink when we are "at the money"

  where f is a test function ranges over "smooth" functions from $\mathbb{R}^d$ to $\mathbb{R}$.

- With a weak error criterion, only the distribution of $\hat{S}_{mh}$ matters.
- In finance applications we generally care about derivatives prices and so the weak criterion is more appropriate.

EFMD EQUIS ACCREDITED AACSB ACCREDITED ASSOCIATION AMBA ACCREDITED

The Euler Scheme VI

Weak and Strong Order of Convergence

Given an error criterion, we can assess the performance of a scheme via its order of convergence:

- We say the discretization $\hat{S}$ has a strong order of convergence of $\beta > 0$ if

$$\mathbb{E}\left(\left\|\hat{S}_{mh} - S_T\right\|\right) \leq ch^{\beta},$$

  for some constant c and sufficiently small h.
- We say the discretization $\hat{S}$ has a weak order of convergence of $\beta > 0$ if

$$\left|\mathbb{E}\left(f(\hat{S}_{mh}) - f(S_T)\right)\right| \leq ch^{\beta},$$

  for some constant c (possibly depending on f), all sufficiently small h, and all sufficiently smooth f .

The Euler Scheme VII

- A larger value of β is better.
- In practice, often the case that a given discretization scheme will have a smaller strong order of convergence than its weak order of convergence.
  Example: The Euler scheme has a strong order of $\beta = 1/2$ but its weak order is $\beta = 1$
- The conditions on f in weak order definition may not met in practice.
  Example: If f represents the payoff of a simple European call option, then f will not be differentiable and so f not sufficiently smooth.
- As a result, experimentation is often required to understand which schemes perform better for a given payoff f and / or SDE S_t .

Outline

- The Euler Scheme
- The Milstein Scheme
- Improvements and Extensions

Milstein scheme I

WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

- The Milstein scheme is based on a higher order Taylor expansion.
- The idea is to apply Ito's Lemma to $b(S_t)$ to construct a better approximation for the diffusion term over the interval $[(k-1)h, kh]$.

Milstein approximation

Suppose we have an SDE $dS_t = a(S_t)dt + b(S_t)dW_t$

$$\hat{S}_{kh} \approx \hat{S}_{(k-1)h} + a\left(\hat{S}_{(k-1)h}\right) + b\left(\hat{S}_{(k-1)h}\right)\sqrt{h}Z_k$$
$$+ \frac{1}{2}b'(\hat{S}_{(k-1)h})b(\hat{S}_{(k-1)h})h(Z_k^2 - 1),$$

where Z_ks are IID $N(0,1)$.

- Under some smoothness conditions it can be shown that the Milstein scheme has a weak and strong order of convergence of $\beta = 1$.

Outline

- The Euler Scheme
- The Milstein Scheme
- Improvements and Extensions

Change of Variables

- Given a scheme, we can choose which process we apply to.
- We can apply our scheme to S_t or to $Y_t := g(S_t)$ where g is a smooth invertible function.
- If we apply it to Y_t then $\hat{S}_{kh} := g^{-1}(\hat{Y}_{kh})$ is the corresponding discretized scheme for S_t .
- Most of the time a particular transformation seems intuitive. For example, if S_t represent a stock price then it makes sense (why?) to apply the scheme to $Y_t := \log(S_t)$ with $g^{-1}(\hat{Y}_{kh}) = \exp(\hat{S}_{kh})$.
- An important advantage of this idea is that we can seek a g with a view to minimizing discretization error.
- A common strategy is to choose a g (when possible) such that the dynamics of Y_t have a constant diffusion coefficient.

Simulation of Jump-Diffusion Processes-I

- Consider a jump-diffusion process of the form

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t + c(X_{t-}, Y_{N_{t-}+1})dN_t,$$

where N_t is a Poisson process (independent of W_t) with parameter λ.

limit from the left

- The Y_i 's are IID random variables independent of W_t . Note $X_{t-} := \lim u \uparrow tX_u$ so if t is a jump time then X_{t-} is the value of the process immediately before t.
- If the nth jump in the Poisson process occurs at time t, then $X_t - X_{t-} = c(X_{t-}, Y_n)$. If a jump does not occur at time t then $X_{t-} = X_t$.

Simulation of Jump-Diffusion Processes-II

A natural strategy to simulate the jump process on $[0, T]$ is

Algorithm

1. First simulate the arrival times in the Poisson process up to time T.
2. Use a pure diffusion discretization between the jump times.
3. At the nth jump time τ_n, simulate the jump size $c(\hat{X}_{\tau_n -}, Y_n)$ conditional on the value of the discretized process $\hat{X}_{\tau_n -}$, immediately before τ_n.

Sample paths of short-term interest rates

Let r_t be the risk-free interest rate applying to the time interval $(t, t + dt)$. This may be called the instantaneous interest rate, although it is often referred to as the short rate. There are different models for the short rate. We will cover:

- The Vasicek model, characterized by a stochastic differential equation featuring mean reversion:

$$dr_t = \gamma(\bar{r} - r_t)dt + \sigma dW_t$$

- The Cox-Ingersoll-Ross (CIR) model, which is quite similar to the Vasicek model, but involves a slight change in the volatility term:

$$dr_t = \gamma(\bar{r} - r_t)dt + \sqrt{\alpha r_t} dWt.$$

Vasicek Model

- Vasicek model implies short rate dynamics following an Ornstein-Uhlenbeck process.
- r_t can get negative.
- It is a Gaussian and mean-reverting process.
- In order to get the solution, we can apply Ito's lemma to the process $f(r_t) = r_t e^{\gamma t}$.
- This implies

$$\mathbb{E}(r_t) = r_0 e^{-\gamma t} + \bar{r}(1 - e^{-\gamma t})$$

$$Var(r_t) = \frac{\sigma^2}{2\gamma}(1 - e^{-2\gamma t})$$

- To generate sample paths (exact simulation) with time step δt we can use,

$$r_{t+\delta t} = r_t e^{-\gamma \delta t} + \bar{r}(1 - e^{-\gamma \delta t}) + \sigma \sqrt{\frac{(1 - e^{-2\gamma \delta t})}{2\gamma}} Z,$$

where $Z \sim N(0, 1)$.

CIR Model

- It also has the mean-reverting property.
- Short rate stays positive as the transition law from r_0 to r_t is represented in terms of χ^2 distribution:

$$r_t = \frac{\alpha(1 - e^{-\gamma t})}{4\gamma}\chi^2(\nu)$$

with degrees of freedom $4\bar{r}\gamma/\alpha$ and non-centrality parameter $\nu = \frac{4\gamma e^{-\gamma t}}{\alpha(1-e^{-\gamma t})} r_0$.

- To generate sample paths via Euler scheme with time step δt we can use:

$$r_{t+\delta t} = \gamma\bar{r}\delta t + (1 - \gamma\delta t)r_t + \sqrt{\alpha r_t \delta t}Z,$$

where $Z \sim N(0, 1)$.

- We can also generate exact sample paths by using the known distribution of r_t.