Course Notes on Computational Finance

1

Modelling Financial Options

1. Modeling of Financial Options

1.1 Options

Definition (Option)

An *option* is the right (but not the obligation) to buy or sell a risky asset at a prespecified fixed "strike" price K until a maturity time T.

The terms of the option contract are fixed by the *writer*. The *holder* of the option pays a premium V for its purchase.

Exercising the option means to buy or sell the *underlying* asset for the price K according to the option's contract. An option with the right to buy the underlying is called *call*, and the option to sell is called *put*.

Question: What is the fair premium V?

This depends on the price K, on the price S_0, on T, and on market data such as the rate r or the volatility σ.

The *volatility* σ measures the size of fluctuations of the asset price S_t, and hence indicates the risk.

A *European* option can only be exercised at maturity ($t = T$); an American option can be exercised anytime during the life time $0 \leq t \leq T$.

The value of the premium V at maturity is easy to assess: it is the *payoff*.

1. Call in $t = T$

The holder of the option has two alternatives to acquire the asset:

(a) She buys it on the spot market and pays S_T, or

(b) exercises the call option and pays the strike price K.

The rational holder optimizes her position.

1st case: $S_T \leq K \Rightarrow$ The holder pays S_T on the spot market, and lets the option expire. Then the option is worthless, $V = 0$.

2nd case: $S_T > K \Rightarrow$ The holder exercises the call and pays K. And immediately she sells the asset for the spot price S_T. The profit is $S_T - K$, hence $V = S_T - K$.

In summary, the payoff of a call is

$$V(S_T, T) = \begin{cases} 0 & \text{in case } S_T \le K \\ S_T - K & \text{in case } S_T > K \end{cases}$$

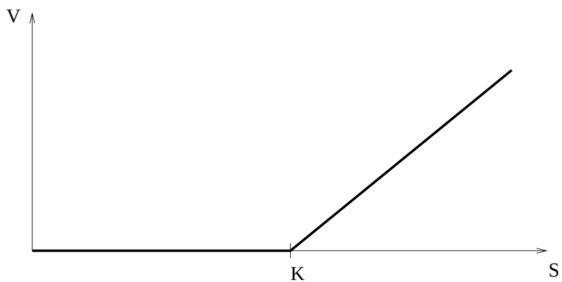
$$= \max\{S_T - K, 0\} \quad =: (S_T - K)^+$$

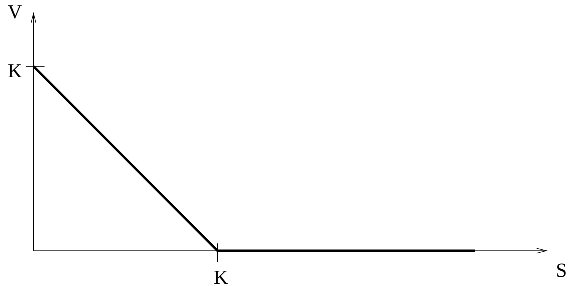
2. Put in $t = T$

Analogous reasoning leads to the payoff of a put:

$$V(S_T, T) = \begin{cases} K - S_T & \text{in case } S_T \le K \\ 0 & \text{in case } S_T > K \end{cases}$$

$$= \max\{K - S_T, 0\} \quad =: (K - S_T)^+$$





Payoff of a call (left) and of a put (right), in $t = T$.

The same arguing is valid for American-style options for any $t \le T$: the payoffs are

$$\begin{aligned} \text{put:} \quad & (K - S_t)^+ \\ \text{call:} \quad & (S_t - K)^+ \end{aligned}$$

The value V for $t < T$, in particular for $t = 0$, is more difficult to determine. The *no-arbitrage-principle* plays a central role. This mere principle leads to **bounds** for V. We give some examples.

The value $V(S,t)$ of an American option can not be smaller than the payoff, because (proof for a put; call is analogous):

> Obviously $V \ge 0$ for all S. *Assume*: $S < K$ and $0 \le V < K - S$. Establish arbitrage as follows: Buy the asset $(-S)$ and the put $(-V)$, and exercise immediately: $(+K)$. By $K > S + V$ this is a risk-free profit $K - S - V > 0$, which contradicts the no-arbitrage-principle.

Hence

$$V^{\text{Am}}_{\text{Put}}(S,t) \ge (K - S)^+ \quad \forall S, t\,.$$

Analogously:

$$V^{\text{Am}}_{\text{Call}}(S,t) \ge (S - K)^+ \quad \forall S, t\,.$$

Also the inequality

$$V^{\mathrm{Am}} \geq V^{\mathrm{Eu}}$$

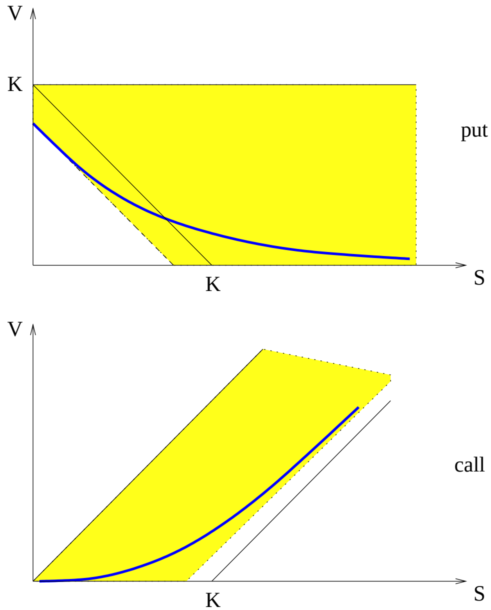
holds since an American option embraces the European option. When no dividend is paid, the put-call parity

$$S + V_{\mathrm{Put}} = V_{\mathrm{Call}} + K\mathrm{e}^{-r(T-t)}$$

holds for European-style options. This leads to further bounds, for example, to

$$V^{\mathrm{Eu}}_{\mathrm{Put}} \geq K\mathrm{e}^{-r(T-t)} - S.$$

The figure illustrates the a-priori bounds for European options on assets that pay no dividends for $0 \leq t \leq T$ (for $r > 0$).



Definition (historic volatility)

The historic volatility σ is the standard deviation of S_t. It is scaled by $\frac{1}{\sqrt{\Delta t}}$ since the data are returns sampled at Δt. In reality, σ is not constant, but the classic Black–Scholes-model takes it as constant. The empirical determination of market parameters (such as σ) is an ambitious task (*calibration*).

Notice that each option involves three prices, namely, the price S_t of the underlying asset, the strike price K and the premium V of the option.

Definition

Options with the above payoffs $\Psi(S) := (K-S)^+$ or $\Psi(S) := (S-K)^+$ on a single asset are called *standard options*, or *vanilla options*. There are many other kinds of options with other features. These other types of options are called *exotic*.

Examples of exotic options

Basket: The underlying is a basket of several assets, e.g., $\sum_{i=1}^{m} w_i S_i(t)$, where S_i is the market price of the ith asset, $m > 1$.

Options with *other payoffs*, such as the *binary* put with

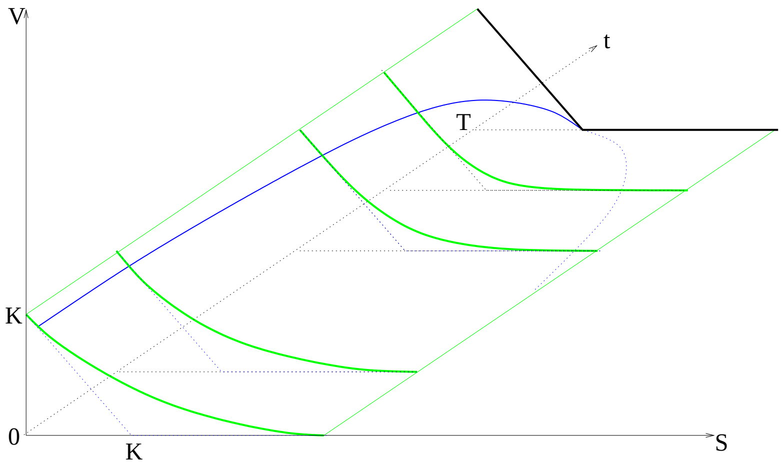
$$\text{payoff} = \begin{cases} 0 & \text{in case } S_t > K \\ 1 & \text{in case } S_t \leq K\,. \end{cases}$$

Path dependence: For instance, the payoff $(\frac{1}{T}\int_0^T S(t)\,\mathrm{d}t - K)^+$ involves the average value, which depends on the path of $S(t)$ (*average price call*).

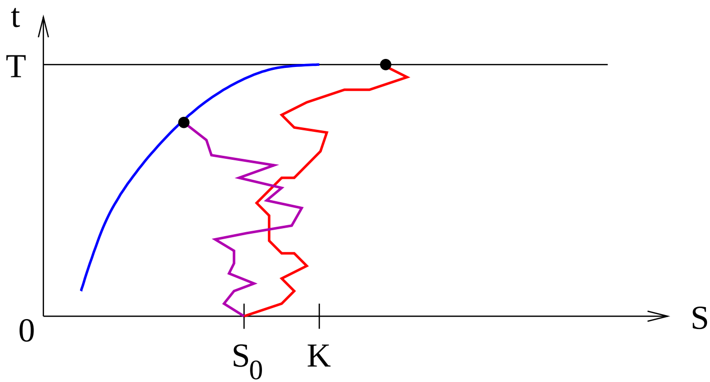
Barrier: For instance, an option ceases to exist when S_t reaches a prespecified barrier B.

On the Geometry of options

The values $V(S,t)$ obey the bounds sketched above, see the illustration of an American put. $V(S,t)$ can be interpreted as surface (figure: in green) over the half strip $0 \leq t \leq T$, $S > 0$. This $V(S,t)$ is called *value function*. At the *early-exercise* curve, the surface merges in the plane defined by the payoff.



Importance: When the market price S_t reaches this (blue) curve, immediate exercise is optimal: invest K for the interest rate r. The situation is sketched in an (S,t)-plane for an American put that pays no dividend.



For American call options *with* dividend payment the situation is analogous. The geometry at early-exercise curves will be discussed in Chapter 4. The curve must be calculated numerically.

1.2 Mathematical Model

A. Black–Scholes Market

Here we discuss mathematical models of how paths S_t may behave. We list some assumptions, which essentially go back to Black, Scholes and Merton (1973, Nobel-Prize 1997). These classic assumptions lead to a partial differential equation (PDE), the famous Black–Scholes equation:

$$\boxed{\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + r S \frac{\partial V}{\partial S} - r V = 0}$$

This equation is a symbol representing the classic theory. Each solution $V(S,t)$ of a European standard option must solve this PDE, satisfying for $t = T$ the terminal condition $V(S,T) = \Psi(S)$ where Ψ denotes the payoff.

Assumptions of the Model

1. There is no arbitrage.
2. The market is frictionless. That is, there are no transaction costs, and rates for lending and borrowing money are equal. All variables are perfectly divisible ($\in \mathbb{R}$). And individual trading does affect the market price.
3. The price S_t follows a geometric Brownian motion (explained later).
4. Technical assumptions:

   r and σ are constant for $0 \leq t \leq T$. No dividends are paid in $0 \leq t \leq T$.

Provided these assumptions hold (some can be weakened), **the value function of a European standard option solves the Black–Scholes equation**. Hence a possible approach to price a European option is to solve the Black–Scholes equation. There is an analytical solution; this is given at the end of this chapter (with δ a continuous dividend rate).

The above model of a finance market is the classic approach; there are other market models.

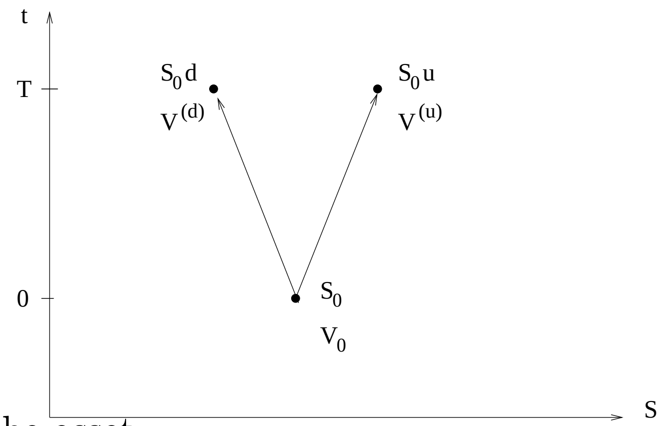
The model with its geometric Brownian motion (Section 1.5) is a continuous-time model, $t \in \mathbb{R}$. There are also discrete-time models, which consider only discrete time instances. Other market models do not use the geometric Brownian motion. Such models are mainly working with jump processes.

Numerical Tasks:

- Computation of $V(S,t)$, in particular for $t = 0$, with early-exercise curve for American options,
- Computation of sensitivities ("Greeks"), such as $\frac{\partial V(S,0)}{\partial S}$,
- Calibration, which means to estimate parameters that match empirical data.

B. Risk-Neutral Probabilities (One-Period Model)

Assumptions: $0 < d < u$, and the situation of the figure below. There are only two time instances: $0, T$, and two possible future asset prices $S_0 d$, $S_0 u$. V_0 denotes the (unknown) value of the option "today" for $t = 0$, and S_0 is the current value of the asset.



Consider a portfolio with two positions:

1. Δ shares of the asset
2. a short position of one option written on the asset

With Π_t denoting the wealth function, the value Π_0 of the portfolio at the time 0 is

$$\Pi_0 = S_0\,\Delta - V_0\,.$$

The number Δ is to be determined. At time T the value of the underlying is "up" or "down" and the portfolio is

$$\Pi^{(u)} = S_0 u\Delta - V^{(u)}$$
$$\Pi^{(d)} = S_0 d\Delta - V^{(d)}\,.$$

$V^{(u)}$ and $V^{(d)}$ are fixed by the payoff. Choose Δ such that the portfolio becomes riskless at time T. That is, the value of the portfolio should be the same, no matter whether the market price goes "up" or "down",

$$\Pi^{(u)} = \Pi^{(d)} =: \Pi_T .$$

Consequently,

$$S_0 \Delta (u-d) = V^{(u)} - V^{(d)}$$

$$\text{or} \qquad \Delta = \frac{V^{(u)} - V^{(d)}}{S_0 u - S_0 d} .$$

With this special value of Δ the portfolio is riskless. Invoking the no-arbitrage principle, we conclude: Any other risk-free investment must have the same value, because otherwise arbitrageurs would make a riskless profit by exchanging the investments. Hence: $\Pi_T = \Pi_0 \, \mathrm{e}^{rT}$

An elementary calculation shows

$$V_0 = \mathrm{e}^{-rT} \left(V^{(u)} q + V^{(d)} (1-q) \right)$$

with $q := \frac{\mathrm{e}^{rT} - d}{u - d}$. This formula has the structure of an expectation. In case $0 < q < 1$ (this requires $d < \mathrm{e}^{rT} < u$, a condition guaranteeing absence of arbitrage*), then this q induces a probability Q, and

---

* What are the arbitrage strategies in case $d \geq \mathrm{e}^{rT}$ or $\mathrm{e}^{rT} \geq u$?

$$\boxed{V_0 = \mathrm{e}^{-rT}\, \mathsf{E}_{\mathsf{Q}}[V_T]}$$

[Recall that in a discrete probability space with a probability P

$$\mathsf{E}_{\mathsf{P}}[X] = \sum_{i=1}^{n} x_i \; \mathsf{P}(X = x_i)$$

holds, where X is a random variable.] The special probability Q defined above is called **risk-neutral probability**. For S_0 we have

$$\mathsf{E}_{\mathsf{Q}}[S_T] = \underbrace{\frac{\mathrm{e}^{rT} - d}{u - d}}_{=q}\; S_0 u + \underbrace{\frac{u - \mathrm{e}^{rT}}{u - d}}_{=1-q}\; S_0 d = S_0 \mathrm{e}^{rT},$$

or

$$\boxed{S_0 = \mathrm{e}^{-rT}\, \mathsf{E}_{\mathsf{Q}}[S_T]\,.}$$

Summary:

In case the portfolio is risk-free (achieved by the above special value of Δ) and when $0 < q < 1$ with $q = \frac{e^{rT}-d}{u-d}$, then there is a probability Q, such that

$$V_0 = e^{-rT}\mathsf{E}_{\mathsf{Q}}[V_T] \quad \text{and}$$
$$S_0 = e^{-rT}\mathsf{E}_{\mathsf{Q}}[S_T]\,.$$

The quantity Δ is called *Delta*. Later we shall see $\Delta = \frac{\partial V}{\partial S}$ in the time-continuous situation. This is the first and most important example of the "Greeks", others are $\frac{\partial^2 V}{\partial S^2}, \frac{\partial V}{\partial \sigma}, \ldots$

Δ is the key for "*Delta-Hedging*", for minimizing or eliminating the risk of the writer of an option.

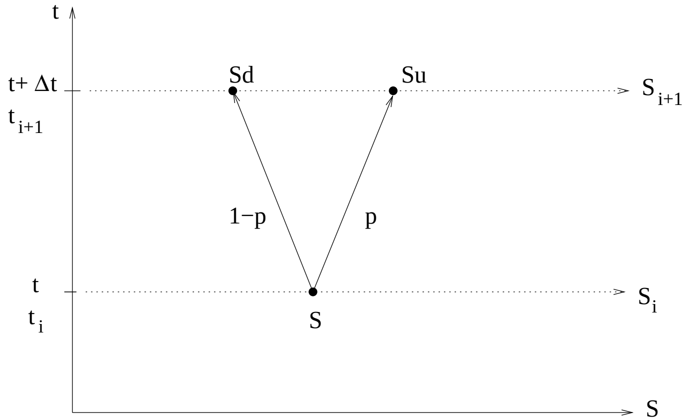
Remark: The relation

$$e^{-rT}\ \mathsf{E}_{\mathsf{Q}}[S_T] = S_0$$

for all T is the *martingale property* of the discounted process $e^{-rt}S_t$ with respect to the probability Q.

1.3 Binomial Method

For the numerical pricing of options, the continuous time must be discretized. Among the many possible approaches the tree methods have the reputation to be both simple and robust. The simplest version uses a binomial tree. The Black–Scholes model results in the limit when the fineness of the binomial tree goes to zero.



Define an equidistant time discretization:

M: number of time steps
$\Delta t := \frac{T}{M}$
$t_i := i \cdot \Delta t, \quad i = 0, ..., M$
$S_i := S(t_i)$

On the S_i-axes we shall define discrete $S_{j,i}$-values.

Assumptions

(Bi1) The market price over one period Δt can only take two values, Su or Sd with $0 < d < u$.

(Bi2) Let the probability of an "up" motion be p, $\mathsf{P}(\text{up}) = p$, with $0 < p < 1$.

(Bi3) Expectation and variance equal those of the continuous-time model (for geometric Brownian motion S_t with riskless growth rate r).

(Bi1) and (Bi2) define the framework of a binomial process with probability. The free parameters u, d, p are to be determined such that (Bi1) – (Bi3) hold.

Remarks

1. It turns out that P is the risk-neutral probability Q. Literature on the stochastic background: [Musiela&Rutkowski: Martingale Methods in Financial modeling], [Shreve: Stochastic Calculus for Finance II (Continuous-time models)].
2. In Section 1.5D we shall show for the continuous-time Black–Scholes model

$$\mathsf{E}[S_t] = S_0 \mathrm{e}^{r(t-t_0)}$$
$$\mathsf{E}[S_t^2] = S_0^2 \mathrm{e}^{(2r+\sigma^2)(t-t_0)}$$

Set S_i for S_0, S_{i+1} for S_t and Δt for $t - t_0$.

3. The expectations are conditional expectations since the initial values $S(t_0)$ or S_i are given.

Conclusion for the step $i \longrightarrow i+1$:

$$\mathsf{E}[S(t_{i+1}) \mid S(t_i) = S_i] = S_i \mathrm{e}^{r\Delta t}$$
$$\mathrm{Var}[S(t_{i+1}) \mid S(t_i) = S_i] = S_i^2 \mathrm{e}^{2r\Delta t}(\mathrm{e}^{\sigma^2 \Delta t} - 1)$$

The expectation of the discrete model is

$$\mathsf{E}[S_{i+1}] = p\ S_i u + (1-p)\ S_i d.$$

Equating with the expression of the continuous-time model shows

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

This is the first equation for the three unknowns u, d, p. This gives

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

For $0 < p < 1$ we require

$$d < \mathrm{e}^{r\Delta t} < u\,.$$

This must hold because otherwise arbitrage is possible. (Compare with Section 1.2B to see that p is the q and represents the risk-neutral probability.)

Equating variances leads to

$$\begin{aligned}\mathsf{Var}[S_{i+1}] &= \mathsf{E}[S_{i+1}^2] - (\mathsf{E}[S_{i+1}])^2 \\ &= p\ (S_i u)^2 + (1-p)\ (S_i d)^2 - S_i^2 (pu + (1-p)d)^2 \\ &\stackrel{!}{=} S_i^2 \mathrm{e}^{2r\Delta t}(\mathrm{e}^{\sigma^2 \Delta t} - 1)\,,\end{aligned}$$

which amounts to

$$\mathrm{e}^{2r\Delta t + \sigma^2 \Delta t} = pu^2 + (1-p)d^2 .$$

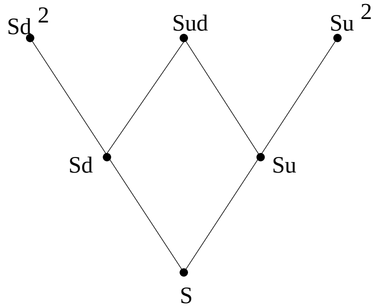
A third equation can be posed arbitrarily. For example, a kind of symmetry is expressed by

$$u \cdot d = 1.$$

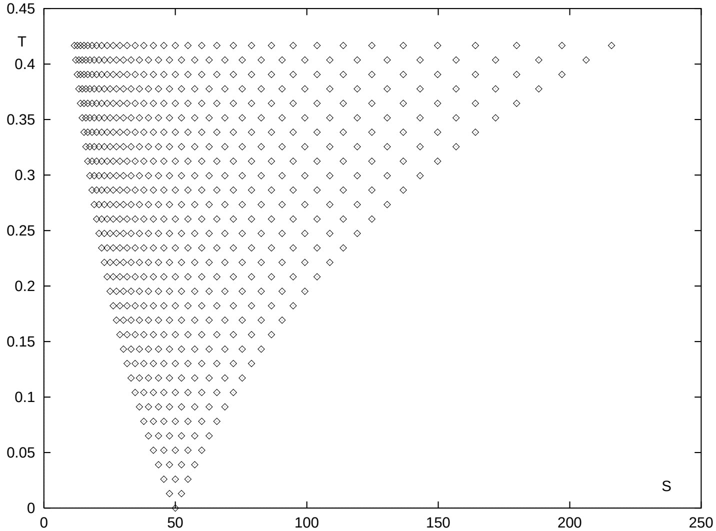
The resulting system of nonlinear equations for u, d, p is

$$\begin{aligned}
\beta :&= \frac{1}{2}(\mathrm{e}^{-r\Delta t} + \mathrm{e}^{(r+\sigma^2)\Delta t}) \\
u &= \beta + \sqrt{\beta^2 - 1} \\
d &= 1/u = \beta - \sqrt{\beta^2 - 1} \\
p &= \frac{\mathrm{e}^{r\Delta t} - d}{u - d}
\end{aligned}$$

This defines the grid of a tree. By the requirement $ud = 1$, this simple tree is rigid in the sense that its parameters u, d, p do not depend on K or S_0. The tree is recombining.



Since $S_{i+1} = \alpha S_i$, $\alpha \in \{u, d\}$ the "branches" of the tree grow exponentially.



The S-values of the grid are

$$S_{j,i} := S_0 u^j d^{i-j}, \quad j = 0, ..., i, \ i = 1, ..., M.$$

Valuation on the Tree

For t_M the value of the option is known from the payoff $\Psi(S) = (S-K)^+$ or $(K-S)^+$:

$$V_{j,M} := \Psi(S_{j,M})$$

By he risk-neutral evaluation principle (Section 1.2B),

$$V_i = \mathrm{e}^{-r\Delta t}\,\mathsf{E}[V_{i+1}],$$

or applied to the tree:

$$\boxed{V_{j,i} = \mathrm{e}^{-r\Delta t} \cdot \left(pV_{j+1,i+1} + (1-p)V_{j,i+1}\right).}$$

This relation establishes a recursion, which starts with $i = M-1$ and prices V at the nodes, until $V_0 := V_{0,0}$.

In case of an **American option**, each node requires a check whether early exercise is reasonable. The holder of the option optimizes her position by comparing the payoff $\Psi(S)$ with the *continuation value*: she chooses the larger value. This requires to modify the above recursion. We denote the continuation value

$$V_{j,i}^{\text{cont}} := \mathrm{e}^{-r\Delta t}\,(pV_{j+1,i+1} + (1-p)V_{j,i+1}).$$

For European options $V_{j,i} := V_{j,i}^{\text{cont}}$.

For American options $V_{j,i} := \max\{\Psi(S_{j,i}),\, V_{j,i}^{\text{cont}}\}$, or

call: $V_{j,i} := \max\{(S_{j,i} - K)^+, V_{j,i}^{\text{cont}}\}$

put: $V_{j,i} := \max\{(K - S_{j,i})^+, V_{j,i}^{\text{cont}}\}$

(principle of dynamic programming)

The two different decisions, either *holding* or *exercising* the American-style option, have a geometrical aspect: In the (S,t)-plane the nodes with $V_{j,i}^{\text{cont}} > \Psi(S_{j,i})$ characterize the *continuation area*, and the other nodes are in the *stopping area*. How the early-exercise curve separates the two areas will be discussed in Chapter 4.

Algorithm (Binomial Method, basic version)

Input: $r,\ \sigma,\ S = S_0,\ T,\ K,$ put or call,
European or American, M

compute: $\Delta t := T/M,\ u,\ d,\ p$ as defined above

$S_{0,0} := S_0$

$S_{j,M} = S_{0,0} u^j d^{M-j},\ j = 0, 1, ..., M$

(for American options in addition $S_{j,i} = S_{0,0} u^j d^{i-j}$
for $0 < i < M$, $j = 0, 1, ..., i$)

$V_{j,M}$ from the payoff

$V_{j,i}$ for $i < M$ by the proper formula

Output: $V_{0,0}$ is approximation for $V(S_0, 0)$

Advantages of the Method

- easy to implement,
- robust, and
- can be adapted to other types of options.

Disadvantages of the Method

– accuracy is rather poor:

  error $O(1/M) = O(\Delta t)$, which is linear convergence. (But the accuracy matches practical requirements.)

– In case V_0 is needed for several values of S, the algorithm must be restarted.

Enhancements

– To avoid oscillations, generalize $ud = 1$ to $ud = \gamma$ and choose γ such that for $t = T$ one node of the tree falls on the strike value K. Then the parameters depend on K and S_0, resulting in a more flexible tree and improved accuracy.

– Discrete dividend payment at time t_D: Cut the tree at t_D and shift the S-values by $-D$. As result, evaluate the tree at $\tilde{S}_0 := S_0 - D\mathrm{e}^{-rt_D}$. (Illustrations in Topic 1 and 5 in the *Topics for CF*.)

– Sensitivities ("greeks") are calculated by difference quotients.

Problems

In the higher-dimensional case (e.g. basket option with three or more assets) it is not obvious how to generalize the tree.
In the literature the above method is often called Cox-Ross-Rubinstein method (CRR). Other extensions: trinomial method; "implied grid" for variable $\sigma(S, t)$.

1.4 Stochastic Processes

This section introduces continuous-time models as they are used by Black, Scholes and Merton. Essentially we discuss (geometric) Brownian motion.

History

Brown (1827): studied erratic motion of pollen.

Bachelier (1900): applied Brownian motion to model asset prices.

Einstein (1905): molecular motion

Wiener (1923): mathematical model

since 1940: Itô and others

Definition (Stochastic Process)

A *stochastic process* is a family of random variables X_t for $t \geq 0$ or $0 \leq t \leq T$.

Each sample results in a function X_t called *path* or *trajectory*.

Definition (Wiener process / standard Brownian motion)

W_t (notation also $W(t)$ or W or $\{W_t\}_{t\geq 0}$) has the properties:

(a) W_t is a *continuous* stochastic process

(b) $W_0 = 0$

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

(d) All increments $\Delta W_t := W_{t+\Delta t} - W_t$ (Δt arbitrary) on non-overlapping t-intervals are *independent*.

(c) means: W_t is distributed normally with $\mathsf{E}[W_t] = 0$ and $\text{Var}[W_t] = \mathsf{E}[W_t^2] = t$.

Remarks

1) “standard”, because it is scalar, driftless, and $W_0 = 0$.
$X_t = a + \mu t + W_t$ with $a, \mu \in \mathbb{R}$ is the general Brownian motion (with drift μ).

2) Consequences (also for $W_0 = a$):

$$\mathsf{E}[W_t - W_s] = 0\ ,\ \text{Var}[W_t - W_s] = t - s \quad \text{for } t > s\,.$$

(show this as exercise)

3) W_t is nowhere differentiable! Motivation:

$$\text{Var}\left[\frac{\Delta W_t}{\Delta t}\right] = \frac{1}{(\Delta t)^2}\text{Var}[\Delta W_t] = \left(\frac{1}{\Delta t}\right)^2 \cdot \Delta t = \frac{1}{\Delta t}$$

tends to ∞ for $\Delta t \to 0$.

4) A Wiener process is self-similar in the sense:

$$W_{\beta t} \overset{\mathrm{d}}{=} \sqrt{\beta}\, W_t$$

(both sides obey the same distribution). More general, there are fractal Wiener processes with

$$W_{\beta t} \overset{\mathrm{d}}{=} \beta^H W_t \,,$$

for the standard Wiener process $H = \frac{1}{2}$. H is the *Hurst-exponent*. Mandelbrot postulated that finance models should use fractal processes.

Importance

The Wiener process is “driving force” of basic finance models.

Discretization/Computation

So far we have considered W_t for continuous-time models ($t \in \mathbb{R}$). Now we approximate W by a discretization. Take $\Delta t > 0$ as a fixed time increment.

$$t_j := j \cdot \Delta t \quad \Rightarrow \quad W_{j\Delta t} = \sum_{k=1}^{j} (W_{k\Delta t} - W_{(k-1)\Delta t}) = \sum_{k=1}^{j} \Delta W_k$$

The ΔW_k are independent, and by Remark 2 satisfy

$$\mathsf{E}(\Delta W_k) = 0, \quad \mathrm{Var}(\Delta W_k) = \Delta t.$$

In case Z is a random variable with $Z \sim \mathcal{N}(0,1)$ [Chapter 2], then

$$Z\sqrt{\Delta t} \sim \mathcal{N}(0, \Delta t).$$

Hence

$$Z \cdot \sqrt{\Delta t} \;\text{ for } Z \sim \mathcal{N}(0,1)$$

serves as model for the process of the ΔW_k.

Algorithm (Simulation of a Wiener process)

start: $t_0 = 0$, $W_0 = 0$; choose Δt.

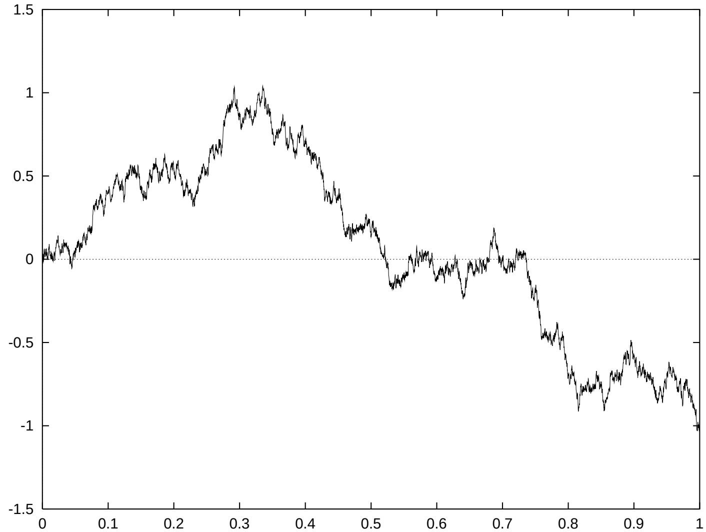
loop $j = 1, 2, ...$:

$t_j = t_{j-1} + \Delta t$

draw $Z \sim \mathcal{N}(0,1)$

$W_j = W_{j-1} + Z\sqrt{\Delta t}$

The W_j denotes a realization of W_t at t_j.



a simulation for $\Delta t = 0.0002$

Stochastic Integral

Motivation:

Assume the price of an asset is described by a Wiener process W_t. Let $b(t)$ be the number of assets in the portfolio at time t. For simplicity assume that there are only discrete trading times

$$0 = t_0 < t_1 < \ldots < t_N = T\,.$$

Hence $b(t)$ is piecewise constant:

$$b(t) = b(t_{j-1}) \quad \text{for} \quad t_{j-1} \le t < t_j \, . \tag{$*$}$$

The resulting trading gain is

$$\sum_{j=1}^{N} b(t_{j-1})(W_{t_j} - W_{t_{j-1}}) \quad \text{for } 0 \le t \le T.$$

Now we approach the time-continuous case and assume arbitrary trading times. The question is whether the sum converges for $N \to \infty$?

For arbitrary b the integral

$$\int_0^T b(t)\, \mathrm{d}W_t$$

does not exist as Riemann–Stieltjes integral. Sufficient for its existence would be a finite *first variation* of W_t.

We show: The first variation $\sum_{j=1}^{N} |W_{t_j} - W_{t_{j-1}}|$ is unbounded.

Proof: Clearly

$$\sum_{j=1}^{N} |W_{t_j} - W_{t_{j-1}}|^2 \le \max_j(|W_{t_j} - W_{t_{j-1}}|) \sum_{j=1}^{N} |W_{t_j} - W_{t_{j-1}}|$$

for any decomposition of the interval $[0, T]$. Now $\Delta t \to 0$. The second variation is bounded, it converges to a $c \neq 0$ (see the Lemma below). By the continuity of W_t, the first factor of the right-hand side goes to 0, and hence the second factor (the first variation) to ∞.

It remains to investigate what happens with the second variation. The relevant type of convergence is *convergence in the mean*,

$$\lim_{N\to\infty} \mathsf{E}[(X - X_N)^2] = 0\,,$$

$$\text{written as:} \quad X = \underset{N\to\infty}{\text{l.i.m.}}\, X_N\,.$$

It remains to show:

Lemma

Denote by $t_0 = t_0^{(N)} < t_1^{(N)} < \ldots < t_N^{(N)} = T$ a sequence of partitions of the interval $t_0 \le t \le T$, with $\delta_N := \max_{j=1}^{N}(t_j^{(N)} - t_{j-1}^{(N)})$. Then:

$$\underset{\delta_N \to 0}{\text{l.i.m.}} \sum_{j=1}^{N} (W_{t_j^{(N)}} - W_{t_{j-1}^{(N)}})^2 = T - t_0$$

Proof: Exercises

Remark: Part of the proof of the lemma comprises the assertions

$$\mathsf{E}[(\Delta W_t)^2 - \Delta t] = 0$$

$$\text{Var}[(\Delta W_t)^2 - \Delta t] = 2 \cdot (\Delta t)^2 .$$

In this probabilistic sense the random variable ΔW_t^2 behaves similarly as Δt. Symbolically this is written

$$\boxed{(\mathrm{d}W_t)^2 = \mathrm{d}t}$$

and will be used for investigations of orders of magnitude.

The construction of an integral for our integrands b

$$\int_{t_0}^{t} b(s)\, \mathrm{d}W_s$$

is based on $\int_{t_0}^{t} b(s)\mathrm{d}W_s := \sum_{j=1}^{N} b(t_{j-1})(W_{t_j} - W_{t_{j-1}})$ for all step functions b in the sense of $(*)$.

For more general b we take step functions converging to b in the mean. For literature see [Øksendal: Stochastic Differential Equations], [Shreve: Stochastic Calculus].

1.5 Stochastic Differential Equations

A. Integral Equation

Definition (Diffusion model)

The integral equation

$$X_t = X_{t_0} + \int_{t_0}^{t} a(X_s, s)\,\mathrm{d}s + \int_{t_0}^{t} b(X_s, s)\,\mathrm{d}W_s$$

for a stochastic process X_t is called Itô stochastic differential equation (SDE). Its symbolic notation is

$$\mathrm{d}X_t = a(X_t, t)\,\mathrm{d}t + b(X_t, t)\,\mathrm{d}W_t$$

Solutions of this stochastic differential equation (that is, of the integral equation) are called *stochastic diffusion*, or *Itô-process*. The term $a(X_t, t)$ is the *drift* term, and $b(X_t, t)$ is the *diffusion*.

Special cases

- The Wiener process is included with $X_t = W_t$, $a = 0$, $b = 1$.
- In the deterministic case $b = 0$ holds, i.e. $\frac{\mathrm{d}X_t}{\mathrm{d}t} = a(X_t, t)$.

Algorithm (analogous as for the Wiener process)

is based on the discrete version

$$\Delta X_t = a(X_t, t)\,\Delta t + b(X_t, t)\,\Delta W_t$$

with increments ΔW and Δt as in Section 1.4. Let y_j denote an approximation of X_{t_j}.

Start: $t_0,\ y_0 = X_0\,;$ choose $\Delta t\,.$

loop: $j = 0, 1, 2, \ldots$

$$t_{j+1} = t_j + \Delta t$$

$$\Delta W = Z\sqrt{\Delta t} \quad \text{with} \quad Z \sim \mathcal{N}(0, 1)$$

$$y_{j+1} = y_j + a(y_j, t_j)\Delta t + b(y_j, t_j)\Delta W$$

Since $\mathrm{d}W^2 = \mathrm{d}t$, we expect an order of only $\frac{1}{2}$; we come back to this in Chapter 3.

B. Application to the Stock Market

Model (GBM = geometric Brownian motion)

$$\boxed{\mathrm{d}S_t = \mu S_t \,\mathrm{d}t + \sigma S_t \,\mathrm{d}W_t}$$

This is an Itô-stochastic SDE with $a = \mu S_t$ and $b = \sigma S_t$. This SDE is *linear* as long as μ and σ do not depend on S_t. For Black and Scholes μ and σ are constant.

(This fills the gap GBM in Assumption 3 in Section 1.2 in the market model.)

μ is interpreted as growth rate, and σ as volatility. The relative change is described by

$$\frac{\mathrm{d}S_t}{S_t} = \mu \,\mathrm{d}t + \sigma \,\mathrm{d}W_t \,.$$

The classic theory of Black, Scholes and Merton (and a significant part of this chapter) assumes a GBM with constant μ, σ.

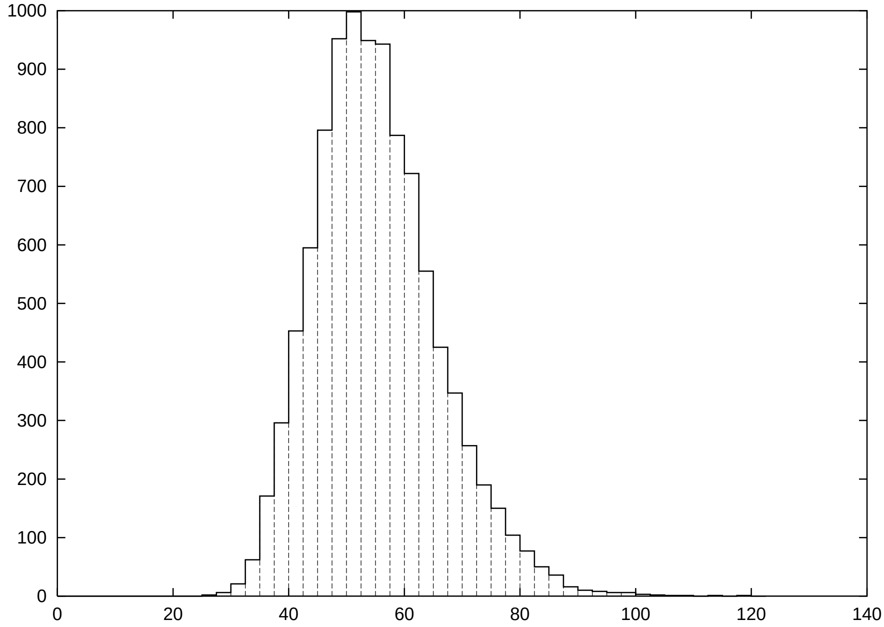
(Bachelier's model was

$$\mathrm{d}S_t = \mu \,\mathrm{d}t + \sigma \,\mathrm{d}W_t \,;$$

here the price S_t can become negative.)

Recommendation

Implement the algorithm (with Z from Chapter 2), and integrate the GBM for a chosen set of parameters (for instance $S_0 = 50$, $\mu = 0.1$, $\sigma = 0.2$) 10000 times until $t = 1$. Then distribute the obtained values S_1 in subintervals, and count the values. This yields a histogram reflecting a lognormal distribution (see figure).



Consequence:

From

$$\frac{\Delta S}{S} = \mu \Delta t + \sigma \, \Delta W$$

we conclude for the distribution of the $\frac{\Delta S}{S}$:

1) distributed normally
2) $\mathsf{E}[\frac{\Delta S}{S}] = \mu \Delta t$
3) $\text{Var}[\frac{\Delta S}{S}] = \sigma^2 \Delta t$

together: $\frac{\Delta S}{S} \sim \mathcal{N}(\mu \Delta t, \sigma^2 \Delta t)$

This offers a way to calculate volatilities σ empirically: For a sequence of trading days collect the data $\frac{\Delta S}{S}$, call them R_i *(returns)*, where R_{i+1} and R_i are measured at time distance Δt. Assuming that GBM is appropriate to describe the returns, σ is obtained as

$$\sigma = \frac{1}{\sqrt{\Delta t}} * \text{ standard deviation of the } R_i \, .$$

This specific value of σ, based on data of the past, is called *historic volatility* (for the *implied* volatility see the Exercises.)

S under GBM can be approximated by the above algorithm as long as $\Delta t > 0$ is small enough, and $S > 0$.

Other models

GBM is continuous, and its density has thin tails, which often fails to describe real asset prices observed in the market. Therefore also other stochastic processes are used, as jump processes, or processes with stochastic volatility. In the following, we stick to the Itô-SDEs, that is to continuous processes driven by Wiener process.

Mean reversion (often used for interest rate models)

Here R denotes an average level of interest rate. Let us investigate the SDE

$$\mathrm{d}r_t = \alpha(R - r_t)\,\mathrm{d}t + \sigma^{\mathrm{r}} r_t^{\beta}\,\mathrm{d}W_t\,, \quad \alpha > 0$$

for a stochastic process r_t. That is,

$$\begin{aligned} a(r_t, t) &= \alpha(R - r_t) \quad \textit{mean reversion}\text{ drift} \\ b(r_t, t) &= \sigma^{\mathrm{r}}\, r_t^{\beta} \end{aligned}$$

with suitable parameters $R, \alpha, \sigma^{\mathrm{r}}, \beta$ (obtained by calibration). This has the effect on the drift:

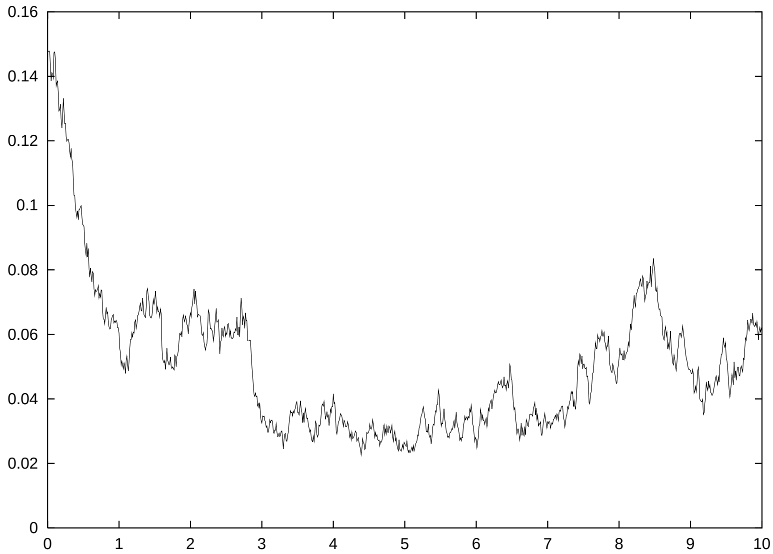
$r_t < R \quad \Rightarrow$ positive growth rate

$r_t > R \quad \Rightarrow$ decay

This effect is superseded by the stochastic fluctuations, but essentially the mean reversion takes care that the order of magnitude of r_t stays close to R, or reverts to R. The parameter α controls the intensity of the reversion.

For $\beta = \frac{1}{2}$, i.e. $b(r_t, t) = \sigma^{\mathrm{r}} \sqrt{r_t}$, the model is called CIR model (Cox-Ingersoll-Ross model).

Figure: A simulation r_t of the Cox-Ingersoll-Ross model for $R = 0.05$, $\alpha = 1$, $\beta = 0.5$, $r_0 = 0.15$, $\sigma^{\mathrm{r}} = 0.1$, $\Delta t = 0.01$



The next extension is to:

Vector-valued processes

Assume $W_t = (W_t^{(1)}, \dots, W_t^{(m)})$ is a m-dimensional Brownian motion. Define for $i = 1, ..., n$

$$X_t^{(i)} = X_{t_0}^{(i)} + \int_{t_0}^{t} a_i(X_s, s)\,\mathrm{d}s + \sum_{k=1}^{m} \int_{t_0}^{t} b_{i,k}(X_s, s)\,\mathrm{d}W_s^{(k)},$$

with vectors

$$X_t = \begin{pmatrix} X_t^{(1)} \\ \vdots \\ X_t^{(n)} \end{pmatrix} \;, \quad a(X_s, s) = \begin{pmatrix} a_1(X_s^{(1)}, \dots, X_s^{(n)}, s) \\ \vdots \\ a_n(X_s^{(1)}, \dots, X_s^{(n)}, s) \end{pmatrix}$$

and matrix

$$\big((b_{i,k})\big)_{i=1,\dots,n}^{k=1,\dots,m}$$

which involves the covariances of the vector process.

Example 1 Heston's model

$$\mathrm{d}S_t = \mu S_t \,\mathrm{d}t + \sqrt{v_t} S_t \,\mathrm{d}W^{(1)}$$

$$\mathrm{d}v_t = \kappa(\theta - v_t)\,\mathrm{d}t + \sigma^{\text{vola}}\sqrt{v_t}\,\mathrm{d}W^{(2)}$$

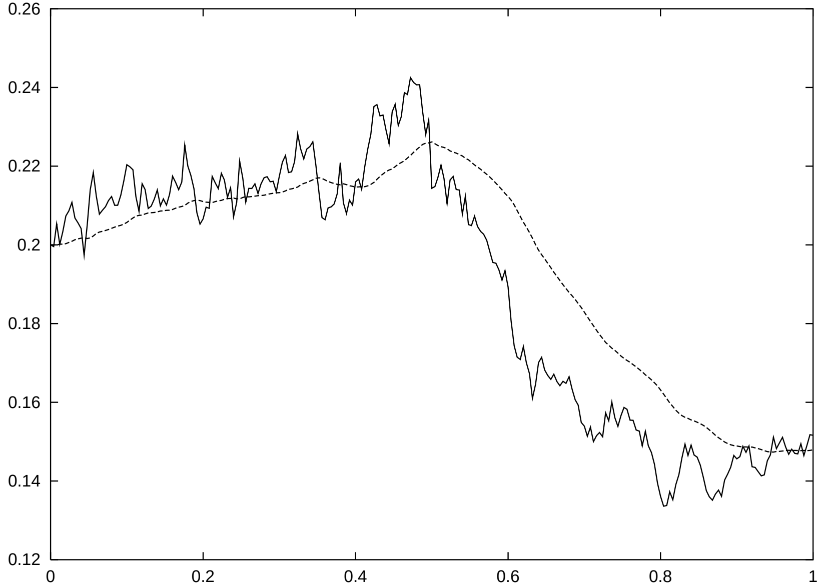
The **stochastic volatility** $\sqrt{v_t}$ is defined via a mean reversion for the variance v_t. This model (with $n = 2$ and $m = 2$) involves parameters $\kappa, \theta, \sigma^{\text{vola}}$, the correlation ρ between $W^{(1)}$ and $W^{(2)}$, an initial value v_0 and a growth rate μ which may be given by a risk-free valuation concept. Altogether, about five parameters must be calibrated. Heston's model is used frequently.

Example 2 volatility tandem

$$dS = \sigma S \, dW^{(1)}$$

$$d\sigma = -(\sigma - \zeta)dt + \alpha\sigma \, dW^{(2)}$$

$$d\zeta = \beta(\sigma - \zeta) \, dt$$



$\alpha = 0.3$, $\beta = 10$; dashed: ζ

Hint: *local* volatility means

$$\sigma = \sigma(t, S_t).$$

C. Itô Lemma

Motivation (deterministic case)

Suppose $x(t)$ is a function, and $y(t) := g(x(t), t)$. The chain rule implies

$$\frac{\mathrm{d}}{\mathrm{d}t} g = \frac{\partial g}{\partial x} \cdot \frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial g}{\partial t}.$$

With $\mathrm{d}x = a(x(t), t)\mathrm{d}t$ this can be written

$$\mathrm{d}g = \left(\frac{\partial g}{\partial x}\, a + \frac{\partial g}{\partial t} \right) \mathrm{d}t$$

Lemma (Itô)

Assume X_t is an Itô process following $\mathrm{d}X_t = a(X_t, t)\,\mathrm{d}t + b(X_t, t)\,\mathrm{d}W_t$ and $g(x, t) \in \mathrm{C}^2$. Then $Y_t := g(X_t, t)$ solves the SDE

$$\mathrm{d}Y_t = \left(\frac{\partial g}{\partial x} a + \frac{\partial g}{\partial t} + \frac{1}{2}\frac{\partial^2 g}{\partial x^2} b^2 \right) \mathrm{d}t + \frac{\partial g}{\partial x} b \; \mathrm{d}W_t \,.$$

That is, Y_t is an Itô process with the same Wiener process as the input process X_t.

Sketch of a proof:

$$\left.\begin{array}{r} t \to t + \Delta t \\ X \to X + \Delta X \end{array}\right\} \to \; g(X + \Delta X, t + \Delta t) = Y + \Delta Y$$

Taylor expansion of g leads to ΔY:

$$\Delta Y = \frac{\partial g}{\partial X} \cdot \Delta X + \frac{\partial g}{\partial t} \Delta t + \text{terms quadratic in } \Delta t, \Delta X$$

Substitute

$$\Delta X = a\,\Delta t + b\,\Delta W$$

$$(\Delta X)^2 = a^2\,\Delta t^2 + b^2 \underbrace{\Delta W^2}_{=O(\Delta t)} + 2ab\,\Delta t\,\Delta W$$

and order the terms according to powers of Δt, ΔW to obtain

$$\Delta Y = \left(\frac{\partial g}{\partial X}\, a + \frac{\partial g}{\partial t} + \frac{1}{2} \frac{\partial^2 g}{\partial X^2}\, b^2 \right) \Delta t + b\, \frac{\partial g}{\partial X}\, \Delta W + \text{t.h.o.}$$

Similar as in Section 1.4, ΔW can be written as sum, and *convergence in the mean* is applied. See [Øksendal].

D. Application to the GBM model

Assume the GBM model

$$\mathrm{d}S = \mu S \,\mathrm{d}t + \sigma S \,\mathrm{d}W$$

with μ and σ constant, i.e. $X = S$, $a = \mu S$, $b = \sigma S$.

1) Let $V(S,t)$ be smooth ($\in C^2$)

$$\Rightarrow \quad \mathrm{d}V = \left(\frac{\partial V}{\partial S}\mu S + \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) \mathrm{d}t + \frac{\partial V}{\partial S}\sigma S \,\mathrm{d}W$$

This is the basic SDE which leads to the PDE of Black and Scholes for the value function $V(S,t)$ of a European standard option.

2) $Y_t := \log(S_t)$, i.e. $g(x) = \log x$

$$\Rightarrow \quad \frac{\partial g}{\partial x} = \frac{1}{x} \quad \text{and} \quad \frac{\partial^2 g}{\partial x^2} = -\frac{1}{x^2}$$

$$\Rightarrow \quad \mathrm{d}\left(\log S_t\right) = \left(\mu - \frac{\sigma^2}{2} \right) \mathrm{d}t + \sigma \,\mathrm{d}W_t$$

Hence the log-prices $Y_t = \log S_t$ satisfy a simple SDE, with the elementary solution:

$$Y_t = Y_{t_0} + \left(\mu - \frac{\sigma^2}{2}\right)(t - t_0) + \sigma(W_t - W_{t_0})$$

$$\Rightarrow \quad \log S_t - \log S_{t_0} = \log \frac{S_t}{S_{t_0}} = \left(\mu - \frac{\sigma^2}{2}\right)(t - t_0) + \sigma(W_t - W_{t_0})$$

$$\Rightarrow \quad S_t = S_{t_0} \cdot \exp\left[\left(\mu - \frac{\sigma^2}{2}\right)(t - t_0) + \sigma(W_t - W_{t_0})\right]$$

For $t_0 = 0$ and $W_{t_0} = W_0 = 0$, this results in

$$\boxed{S_t = S_0 \exp\left[\left(\mu - \tfrac{\sigma^2}{2}\right)t + \sigma W_t\right]}$$

In summary, S_t is exponential function of a Brownian motion with drift.

Implications for $t_0 = 0$:

a) $\log S_t$ is distributed normally

b) $\mathsf{E}[\log S_t] = \mathsf{E}[\log S_0] + (\mu - \frac{\sigma^2}{2})t + 0 = \log S_0 + (\mu - \frac{\sigma^2}{2})t$

c) $\mathsf{Var}[\log S_t] = \mathsf{Var}[\sigma W_t] = \sigma^2 t$

summarizing a) – c) means

$$\log \frac{S_t}{S_0} \sim \mathcal{N}\left((\mu - \frac{\sigma^2}{2})t\,,\,\sigma^2 t\right)$$

d) This leads to the density function of $Y = \log S$

$$\hat{f}(Y) = \hat{f}(\log S_t) = \frac{1}{\sigma\sqrt{2\pi t}}\,\exp\left[-\frac{(\log(S_t/S_0) - (\mu - \sigma^2/2)t)^2}{2\sigma^2 t}\right].$$

And what is the density of S_t? The probabilities of S and Y are the same and hence also the distribution integrals. We apply the transformation theorem (Section 2.2B) for $Y := \log S$ and have the integrands

$$\hat{f}(Y)\,\mathrm{d}Y = \underbrace{\hat{f}(\log S)\frac{1}{S}}_{f(S_t)}\,\mathrm{d}S.$$

Consequently, the density f of the distribution of the asset price S_t is

$$f(S_t, t;\, S_0, \mu, \sigma) := \frac{1}{S_t\sigma\sqrt{2\pi t}}\,\exp\left[-\frac{(\log(S_t/S_0) - (\mu - \sigma^2/2)t)^2}{2\sigma^2 t}\right].$$

This is the density f_{GBM} of the *lognormal distribution*. It describes the probability of the transition $(S_0, 0) \longrightarrow (S_t, t)$ under GBM.

e) Now the last gap in the derivation of the binomial method can be closed: There the continuous model refers to our GBM. As an exercise, realize

$$\mathsf{E}(S) = \int_0^\infty S f(\ldots)\, \mathrm{d}S = S_0\, \mathrm{e}^{\mu(t-t_0)}$$

$$\mathsf{E}(S^2) = \int_0^\infty S^2 f(\ldots)\, \mathrm{d}S = S_0^2\, \mathrm{e}^{(\sigma^2+2\mu)(t-t_0)}$$

1.6 Risk-Neutral Valuation

(This section *sketches* basic ideas and concepts. For a thorough treatment we recommend literature on Stochastic Finance such as [Musiela&Rutkowski: Martingale Methods in Financial modeling])

Recall (from the one-period)

$$V_0 = \mathrm{e}^{-rT} \mathsf{E}_{\mathsf{Q}}[\Psi(S_T)]$$

where Q is the artificial probability of Section 1.2 and $\Psi(S_T)$ denotes the payoff.

For the **model with continuous time** formally the same relation holds. But Q and E_{Q} are different. It turns out that the density of Q is given by $f(S_t, t;\, S_0, r, \sigma)$, with μ replaced by r. Hence the relation

$$V_0 = \mathrm{e}^{-rT} \int_0^\infty \Psi(S_T) \cdot f(S_T, T;\, S_0, r, \sigma)\, \mathrm{d}S_T$$

holds for the GBM-based continuous model. In the following we outline the arguments that lead to this integral.

Fundamental Theorem of Asset Pricing

The market model is free of arbitrage if and only if there is a probability Q such that the discounted asset prices $\mathrm{e}^{-rt}S_t$ are martingales with respect to Q.

Probability space

The same sample space and σ-algebra $(\Omega, \mathcal{F})$ underlying a Wiener process are not specified. The chosen probability P completes $(\Omega, \mathcal{F})$ to the probability space $(\Omega, \mathcal{F}, \mathsf{P})$. The independence of the increments ΔW of the Wiener process depend on P. A process W can be a Wiener process with respect to P, but is no Wiener process with respect to another probability $\hat{\mathsf{P}}$

Martingale

A martingale M_t is a stochastic process with

$$\mathsf{E}[M_t \mid \mathcal{F}_s] = M_s \quad \text{for all } t, s \text{ with } s \leq t\,,$$

where $\mathcal{F}_s$ is a filtration, i.e. a family of σ-algebras with $\mathcal{F}_s \subseteq \mathcal{F}_t \; \forall s \leq t$. A filtration serves as model for the amount of information in a market.

$\mathsf{E}[M_t \mid \mathcal{F}_s]$ is a *conditional expectation*. It can be regarded as expectation of M_t conditional on the amount of information available until time instant s.

M_t martingale means that M_s at time s is the best possible forecast for $t \geq s$.

Martingale with respect to a probability Q: $\mathsf{E}_{\mathsf{Q}}[M_t \mid \mathcal{F}_s] = M_s$ for all t, s with $s \le t$.

Examples of martingales

1) any Wiener process
2) $W_t^2 - t$ for any Wiener process W.
3) A necessary criterion for martingales is the absence of drift.

Essentially, drift-free processes are martingales.

Market Price of Risk

$$\begin{aligned} \mathrm{d}S &= \mu S\,\mathrm{d}t + \sigma S\,\mathrm{d}W \\ &= rS\,\mathrm{d}t + (\mu - r)S\,\mathrm{d}t + \sigma S\,\mathrm{d}W \\ &= rS\,\mathrm{d}t + \sigma S\left[\frac{\mu - r}{\sigma}\,\mathrm{d}t + \,\mathrm{d}W\right] \end{aligned}$$

The investor expects $\mu > r$ as a compensation for the risk, which is represented by σ. $\mu - r$ is the *excess return.*

$$\begin{aligned} \gamma := \frac{\mu - r}{\sigma} &= \text{ “market price of risk”} \\ &= \text{ compensation rate relative to the risk} \end{aligned}$$

Hence

$$dS = rS\,dt + \sigma S\,[\gamma\,dt + dW]\ . \qquad (*)$$

Under the probability P the term in brackets represents a drifted Brownian motion and no (standard) Wiener process.

Girsanov's Theorem

Suppose W is Wiener process with respect to $(\Omega, \mathcal{F}, \mathsf{P})$. In case γ satisfies certain requirements, there is a probability Q such that

$$W_t^\gamma := W_t + \int_0^t \gamma\,ds$$

is a (standard) Wiener process under Q.

(probability theory: Q results from the Theorem of Radon-Nikodym. Q and P are equivalent. For constant γ the requirements of Girsanov are fulfilled.)

Application

Substitute $dW^\gamma = dW + \gamma dt$ in $(*)$ gives

$$dS = rS\,dt + \sigma S\,dW^\gamma .$$

This is a change of drift from μ to r; σ remains unchanged. The path of S_t under the probability Q is defined by the density $f(\ldots, r, \sigma)$. The transition from $f(\ldots, \mu, \sigma)$ to $f(\ldots, r, \sigma)$ amounts to adjusting the probability from P to Q. The discounted $\mathrm{e}^{-rt} S_t$ is drift-free under Q and Martingale. Q is called "risk-neutral" probability.

Trading Strategy

Let X_t be a stochastic vector process of market prices, and b_t denotes the vector with the numbers of shares held in the portfolio. Hence $b_t^{tr} X_t$ is the wealth process of the portfolio.

Example

$$X_t := \begin{pmatrix} S_t \\ B_t \end{pmatrix},$$

where S_t is the market price of the asset underlying an option, and B_t is the value of a risk-free bond.

Notation: V_t is the random variable of the value of an European option.
Assumptions:

(1) There is a strategy b_t *replicating* the payoff of the option at time T,

$$b_T^{tr} X_T = \text{ Payoff } .$$

b_t must be $\mathcal{F}_t$-measurable for all t. (That is, the trader cannot see the future. Note that the value of the payoff is a random variable.)

(2) The portfolio is closed, no money is inserted or withdrawn. This is the *self-financing property* defined as

$$\mathrm{d}(b_t^{tr} X_t) = b_t^{tr} \mathrm{d}X_t \,.$$

(3) The market is free of arbitrage.

$(1), (2), (3) \;\Rightarrow V_t = b_t^{tr} X_t$ for $0 \leq t \leq T$ (otherwise there would exist arbitrage)

We consider a European option and a discounting process Y_t with the property that $Y_t X_t$ is martingale. Then one can show that also $Y_t b_t^{tr} X_t$ is martingale (both with respect to Q).

Implications for European options for $t \leq T$

$$\begin{aligned} Y_t V_t = Y_t b_t^{tr} X_t \quad &= \mathsf{E}_{\mathsf{Q}}[Y_T b_T^{tr} X_T \mid \mathcal{F}_t] \qquad \text{(martingale)} \\ &= \mathsf{E}_{\mathsf{Q}}[Y_T \cdot \text{Payoff} \mid \mathcal{F}_t] \quad \text{(replication)} \end{aligned}$$

When the payoff is a function Ψ of S_T (vanilla-option under GBM), then

$$= \mathsf{E}_{\mathsf{Q}}[Y_T \cdot \Psi(S_T)]$$

(because $W_T - W_t$ is independent of $\mathcal{F}_t$). Discounting with $Y_t = \mathrm{e}^{-rt}$ implies specifically for $t = 0$

$$1 \cdot V_0 = \mathsf{E}_{\mathsf{Q}}[\mathrm{e}^{-rT} \cdot \Psi(S_T)]$$

and hence

$$V_0 = \mathrm{e}^{-rT} \int_0^\infty \Psi(S_T) \cdot f(S_T, T;\, S_0, r, \sigma)\, \mathrm{d}S_T \ .$$

This is called *risk-neutral valuation.*

Literature on Stochastic Finance: [Elliot & Kopp: Mathematics of Financial Markets], [Korn: Option Pricing and Portfolio Optimization], [Musiela & Rutkowski: Martingale Methods in Financial modeling], [Shreve: Stochastic Calculus for Finance].

Outlook

We so far have investigated continuous processes S_t driven by W_t. To compensate for occasional drastic changes in the price of underlying, one resorts to models with stochastic volatility, or to jump processes.

Supplements

The "**Greeks**" mean the sensitivities of $V(S,t;\sigma,r)$ and are defined as

$$\text{Delta} = \frac{\partial V}{\partial S}, \quad \text{gamma} = \frac{\partial^2 V}{\partial S^2}, \quad \text{theta} = \frac{\partial V}{\partial t}, \quad \text{vega} = \frac{\partial V}{\partial \sigma}, \quad \text{rho} = \frac{\partial V}{\partial r}$$

Black–Scholes Formula

For a European call the analytic solution of the Black–Scholes equation is

$$d_1 := \frac{\log\frac{S}{K} + \left(r - \delta + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}$$

$$d_2 := d_1 - \sigma\sqrt{T-t} = \frac{\log\frac{S}{K} + \left(r - \delta - \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}$$

$$V_{\mathrm{C}}(S,t) = S\mathrm{e}^{-\delta(T-t)}F(d_1) - K\mathrm{e}^{-r(T-t)}F(d_2),$$

where F denotes the standard normal cumulative distribution (compare Exercises), and δ is a continuous dividend yield. The value $V_{\mathrm{P}}(S,t)$ of a put is obtained by applying the put-call parity

$$V_{\mathrm{P}} = V_{\mathrm{C}} - S\mathrm{e}^{-\delta(T-t)} + K\mathrm{e}^{-r(T-t)}$$

from which

$$V_{\mathrm{P}} = -S\mathrm{e}^{-\delta(T-t)}F(-d_1) + K\mathrm{e}^{-r(T-t)}F(-d_2)$$

follows.