Angewandte Finanzmathematik 2024: Introduction to the Black–Scholes World

Ari-Pekka Perkkiö

May 2, 2024

Contents

1 Practicalities and background

Upon passing the exam, attending and solving the exercises give a bonus to the final grade.

We assume that the following concepts are familiar:

1. Probability space, random variables, expectation, convergence concepts.
2. Conditional expectations, martingales.
3. The fundamentals of discrete time financial mathematics.

For a remote graphical access to Matlab, you can login to the computers

- math12.math.lmu.de
- mathw0g.math.lmu.de

You will need

1. a program supporting X11-forwarding (e.g. Cygwin),
2. SSH program with rdp connections (e.g. Bitvise),
3. a VPN connection to LRZ (Anyconnect client, downloadable from LMU service portal).

Alternatively, you can use the online version of Matlab.

2 Introduction

2.1 Popular financial products

Throughout the course, S_t^i denotes the market price of an asset i at time t.

Example 2.1 (Put and call options)**.** *A* European call option *on the asset i is a contract where the seller has the obligation to deliver the asset i at the given* maturity time *T for a given* strike price *K. At time T, the buyer has the possibility to exercise the option, that is, to buy the asset from the seller at price K. The gain for the buyer is*

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

since he can get the asset from the option seller at price K and sell it immediately on the market with the market price S_T^i. We call c_C the payoff *of the call option.*

A European put option *on the asset i is a contract where the seller has the obligation to buy the asset i at the given* maturity time T *for a given* strike price K. *At time T, the buyer has the possibility the exercise the option, that is, to sell the asset to the seller at price K. The payoff for the buyer becomes*

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

Put and call options are prototype examples of *Vanilla options* that depend only on the terminal price of the underlying asset. When this is not the case, the option is called *path-dependent*.

Example 2.2 (Asian options). *An* Asian call option *with maturity T and strike K has the payoff*

$$c_{AC} := (\bar{S}_T^i - K)^+,$$

where $\bar{S}_T^i$ is the "average price" of the asset over the time interval $[0,T]$. The exact form of the average price is part of the contract, e.g., it could be arithmetic mean of the prices at given time points $t_1, \ldots, t_N = T$ so that $\bar{S}_T^i = \frac{1}{N}\sum_{k=1}^N S_{t_k}^i$.

For a set A, we denote $\mathbb{1}_A(s) = 1$ if $s \in A$ and $\mathbb{1}_A(s) = 0$ otherwise.

Example 2.3 (Down-and-out and other Barrier options). *Given a strike K, maturity T and a* barrier $B > 0$, *the* down-and-out call option *has the payoff*

$$c_{DOC} := (S_T^i - K)^+ \mathbb{1}_{\mathbb{R}_+}(\min_{t\in[0,T]} S_t^i - B).$$

The payoff of an up-and-in call option *with the same strike and maturity is*

$$c_{UIC} := (S_T^i - K)^+ \mathbb{1}_{\mathbb{R}_+}(\max_{t\in[0,T]} S_t^i - B).$$

Barrier put options have similar payoffs. For example, down-and-in put options *have payoffs of the form*

$$c_{DIP} : (K - S_T^i)^+ \mathbb{1}_{\mathbb{R}_+}(B - \min_{t\in[0,T]} S_t^i).$$

Options that depend on multiple underlying assets are called *rainbow options*.

Example 2.4 (Basket options). *Given a set of assets indexed by $i = 1, \ldots, I$ and positive coefficients a_i, $i = 1, \ldots I$, the payoff of the corresponding* basket call option *is*

$$c_{BC} := \left(\sum_{i=1}^I a_i S_T^i - K\right)^+.$$

Similarly, the basket put option *has the payoff*

$$c_{BP} := \left(K - \sum_{i=1}^I a_i S_T^i\right)^+.$$

Example 2.5 (Spread options)**.** *Given two assets S^1 and S^2, the payoff of the corresponding* spread call option *is*

$$c_{SC} := \left(S_T^1 - S_T^2 - K\right)^+ .$$

Similarly, the spread put option *has the payoff*

$$c_{SP} := \left(K - S_T^1 + S_T^2\right)^+ .$$

Example 2.6 (Calls and puts on max and min)**.** *Given to assets S^1 and S^2, the payoff of the corresponding* call-on-max option *is*

$$c_{MaxC} := \left(\max\{S_T^1, S_T^2\} - K\right)^+ .$$

Similarly, the put-on-min option *has the payoff*

$$c_{MinP} := \left(K - \min\{S_T^1, S_T^2\}\right)^+ .$$

Many options depend on quantities that are not tradable on markets.

Example 2.7 (Options on non-tradables)**.** *Let ξ_T be the temperature (somewhere of interest) at time T, and consider options with the payoffs*

$$(\xi_T - K)^+ \quad \textit{and} \quad (K - \xi_T)^+$$

with a given strike K.

Example 2.8 (American options*)**.** *The holder of an American option may choose to exercise the option at any time before the terminal time T. For example, for an American call on S^i with strike K, the payoff, if the holder chooses to exercises the option at time t, is*

$$(S_t - K)^+.$$

In contrast to all the above options, the holder of an American faces an optimization problem when to exercise the option.

2.2 Exercises

In all the exercises, examples in Matlab online help pages help you to write the actual code.

Exercise 2.2.1. *Write Matlab functions (as .m-files) of the payoff functions in Examples 2.1–2.6. Write them as functions of the underlying asset prices and strikes.*

Exercise 2.2.2. *Using the plot-function, plot the European call option, for a fixed strike K, as a function of the underlying asset price S_T. Plot the European call option as a function of the underlying asset price S_T for two different strikes in the same figure.*

Exercise 2.2.3. *Using the mesh-function (or surf-function), draw a 3D-graph of the spread call option as a function of the underlying asset prices S_T^1 and S_T^2.*

2.3 Basic properties of Brownian motion

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t=0}^T, P)$ be a filtered probability space. We consider continuous time stochastic processes only on the "time interval" $[0,T]$. A family $S := (S_t)_{t\in[0,T]}$ of $\mathbb{R}^d$–valued random variables S_t is called an $\mathbb{R}^d$–valued *continuous time stochastic process*. The process is called *adapted* if S_t is $\mathcal{F}_t$-measurable for each $t \in [0,T]$.

Given $\omega \in \Omega$, the function $t \mapsto S_t(\omega)$ is called as a *path*, or a *trajectory* or a *realization*, of the process S. Instead of considering a stochastic process as an indexed family of $\mathbb{R}^d$-valued random variables, one may thus think of a stochastic process as a family of random paths, trajectories, etc. In some cases (less in this course), it is helpful to think of a stochastic process S as a function $(\omega, t) \mapsto S_t(\omega)$ from the product space $\Omega \times [0,T]$ to $\mathbb{R}^d$. If the paths of a continuous time process are P-almost surely continuous, then the process is called *a continuous stochastic process.*

For a random variable $\eta \in (\Omega, \mathcal{F}, P)$, we denote $\eta \sim N(\mu, \sigma^2)$ when η is a normally distributed random variable with mean μ and standard deviation σ.

Remark 2.9. *We often use the property that for $\eta \sim N(0,\sigma^2)$ and positive integer m, there is a constant L such that $E\eta^{2m} = L\sigma^{2m}$,*

Definition 2.10. *An adapted continuous stochastic process W is a Brownian motion, if it has independent increments in the sense that, for all $0 \le t_0 < t_1 < \cdots < t_n$ the random variables $\{W_{t_i} - W_{t_{i-1}} \mid i = 1, \ldots n\}$ are independent, and $W_t - W_s \sim N(0, t-s)$ for all $0 \le s < t \le T$,*

From now on we assume, unless stated otherwise, that given a Brownian motion W, it starts at zero, that is, $W_0 = 0$.

Exercise 2.3.1. *Show that a Brownian motion W is a martingale, that is, for all $s < t \le T$, $s > 0$, we have $E|W_t| < \infty$ and*

$$E[W_t \mid \mathcal{F}_s] = W_s.$$

Here we assume that the increments of W are independent of the filtration in the sense that, for all $s < t$, the random variable $W_t - W_s$ is indenpendent of $\mathcal{F}_s$. This is the case, .e.g., when the filtration is generated by W.

In the definition of Brownian motion, it possible to omit the assumption that the paths are continuous. This follows from the famous Kolmogorov's continuity criterion. Recall that a continuous function $f : [0,T] \to \mathbb{R}$ is α-Hölder continuous if there is $L \in \mathbb{R}$ such that

$$|f_t - f_s| \le L|t-s|^\alpha \quad \forall\ 0 \le s \le t \le T.$$

Theorem 2.11 (Kolmogorov's continuity criterion)**.** *Let S be a stochastic processes with*

$$E\,|S_t - S_s|^a \le L\,|t-s|^{1+b} \quad \forall\ s < t \tag{2.1}$$

for some constants $a \geq 1$, $b, L > 0$. Then there exists a continuous stochastic process $\tilde{S}$ that is a modification of S in the sense that $P(\tilde{S}_t = S_t) = 1$ for all t. Moreover, $\tilde{S}$ is α-Hölder continuous almost surely for any $\alpha \leq b/a$.

Exercise 2.3.2. *Using Remark 2.9, show that, for any $\epsilon > 0$, Brownian motion has $(1/2 - \epsilon)$-Hölder continuous paths almost surely.*

From the computational perspective, Brownian motion has the important property that it can be approximated by piece-wise constant "discrete-time random walks" that have independent increments. Such random random walks are easy to simulate which is the basis of Monte Carlo methods that is the main topic of the course.

Recall that a sequence of random variables (η^ν) converges in distribution to the random variable η if

$$P(\eta^\nu \leq x) \to P(\eta \leq x)$$

for all $x \in \mathbb{R}$ such that $x \mapsto P(\eta \leq x)$ is continuous (i.e., for all x such that the cumulative distribution function of η is continuous at x). A sequence of vectors of random variables $(\eta_1^\nu, \dots \eta_k^\nu)$ converges in distribution to $(\eta_1, \dots, \eta_k)$ if

$$P((\eta_1^\nu, \dots, \eta_k^\nu) \leq x) \to P((\eta_1, \dots, \eta_k) \leq x)$$

for all $x \in \mathbb{R}^k$ such that $x \mapsto P((\eta_1, \dots, \eta_k) \leq x)$ is continuous.

Theorem 2.12 (The central limit theorem)**.** *Let*

$$\eta^{(n)} = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \xi_k,$$

for an i.i.d. $(\xi_k)_{k=1}^\infty$ sequence of random variables with $E\xi_k = 0$ and $E(\xi_k)^2 = 1$. We have

$$\eta^{(n)} \xrightarrow{d} \eta \tag{2.2}$$

for a random variable $\eta \sim N(0, 1)$.

For continuous time stochastic processes $S^{(n)}$, $n = 1, 2, \dots$ and S, $S^{(n)}$ converges in finite dimensional distributions to S, denoted by

$$S^{(n)} \xrightarrow{fd} S,$$

if, for all integers k and all $0 \leq t_0 < \cdots < t_k \leq T$,

$$(S_{t_0}^{(n)}, \dots, S_{t_k}^{(n)}) \xrightarrow{d} (S_{t_0}, \dots, S_{t_k}).$$

Theorem 2.13. *Let*

$$Y_t^{(n)} := \frac{1}{\sqrt{n}} \sum_{k=1}^{\lfloor nt \rfloor} \xi_k$$

for an i.i.d. $(\xi_k)_{k=1}^{\infty}$ *sequence of random variables with* $E\xi_k = 0$ *and* $E(\xi_k)^2 = 1$. *Then*

$$Y_t^{(n)} \xrightarrow{fd} W$$

for a Brownian motion W.

Proof. Using the central limit theorem and $\frac{\lfloor nt \rfloor}{n} \to t$ when $n \to \infty$, we get

$$Y_t^{(n)} = \frac{\sqrt{\lfloor nt \rfloor}}{\sqrt{n}} \frac{1}{\sqrt{\lfloor nt \rfloor}} \sum_{k=1}^{\lfloor nt \rfloor} \xi_k \xrightarrow{d} \eta \sim N(0,t),$$

as $n \to \infty$. Let now $t < u$. The random variables $Y_u^{(n)} - Y_t^{(n)}$ are independent from the variables $Y_t^{(n)}$, since

$$Y_u^{(n)} - Y_t^{(n)} = \frac{1}{\sqrt{n}} \sum_{k=\lfloor nt \rfloor + 1}^{\lfloor nu \rfloor} \xi_k^{(n)}$$

and the random variables $\xi_k^{(n)}$ are independent. Repeating the previous arguments we get

$$Y_u^{(n)} - Y_t^{(n)} \xrightarrow{d} \eta_{u-t} \sim N(0, u-t).$$

We observe that the variables $\Delta Y_{t_i}^{(n)} := Y_{t_i}^{(n)} - Y_{t_{i-1}}^{(n)}$ are mutually independent for all $0 \le t_0 < t_1 < \cdots < t_N \le T$. Thus the process $Y^{(n)}$ has independent increments, and so

$$P\left(Y_{t_i}^{(n)} - Y_{t_{i-1}}^{(n)} \le x_i, i = 1, \ldots, N\right) = \prod_{i=1}^{N} P\left(Y_{t_i}^{(n)} - Y_{t_{i-1}}^{(n)} \le x_i\right)$$
$$\longrightarrow \prod_{i=1}^{N} \Phi_{0, t_i - t_{i-1}}(x_i) = P(W_{t_i} - W_{t_{i-1}} \le x_i, i = 1, \ldots, N).$$

The proof is finished by the next exercise. □

Exercise 2.3.3. *Recall the continuous mapping theorem: If* $(\eta_0^\nu, \ldots, \eta_k^\nu) \xrightarrow{d} (\eta_0, \ldots, \eta_k)$, *then* $f(\eta_0^\nu, \ldots, \eta_k^\nu) \xrightarrow{d} f(\eta_0, \ldots, \eta_k)$ *for any continuous function* $f : \mathbb{R}^k \to \mathbb{R}^n$.

Use the continuous mapping theorem to finish the proof of Theorem 2.13.

2.4 Exercises

In all the exercises, examples in Matlab online help pages help you to write the actual code. We say that the process defined by

$$B_t := \mu t + \sigma W_t$$

is a Brownian motion with drift μ and volatility σ. Here W is a (standard) Brownian motion.

Exercise 2.4.1. *Write a Matlab function (as an .m-file) that creates a sample path of a Brownian motion with terminal time T, $n+1$ equi-distant discretization points, drift μ and volatility σ. Write it as a function of these parameters and a sample of independent standard normally distributed random variable so that the function maps the sample to a (discretized) sample path of a Brownian motion.*

Exercise 2.4.2. *Plot sample paths of the Brownian motion with different drifts and volatilities in the same figure.*

Download the .mat files from the course page. They contain "classes" consisting of sample paths of a Brownian motion with a given terminal time T.

Exercise 2.4.3. *For paths in "bmpaths.mat", estimate the volatility of each path. Plot the paths in the same figure and label the paths with their volatilities.*

2.5 Quadratic variation

Let

$$D := \bigcup_n D_n,$$

where D_n is the *n-th dyadic partition* of $[0, T]$,

$$D_n := \{i/2^n \in [0, T] \mid i = 0, 1, 2 \ldots, \}.$$

Enumerating $D_n = \{\{t_0^n, t_1^n, \ldots\} \mid t_i^n \leq t_{i+1}^n\}$, we define, for each n, the "discrete quadratic variation" of a stochastic process S by

$$QV_t^n(S) := \sum_{i \geq 1} |S_{t_i^n \wedge t} - S_{t_{i-1}^n \wedge t}|^2,$$

where $s \wedge t := \min\{s, t\}$.

Theorem 2.14. *Let W be a Brownian motion. Then*

$$P(\lim_n QV_t^n(W) = t \quad \forall t) = 1.$$

Proof. Fix $t \in D$. The almost sure convergence $QV_t^n(W) \to t$ is equivalent to the almost sure convergence $\sum_{t_n^i \leq t} Z_i \to 0$ for

$$Z_i := (W_{t_{i+1}^n \wedge t} - W_{t_i^n \wedge t})^2 - 2^{-n}.$$

Note first that $t_{i+1}^n - t_i^n = 2^{-n}$ so that $W_{t_{i+1}^n \wedge t} - W_{t_i^n \wedge t}$ are mutually independent and $N(0, 2^{-n})$-distributed. We have $E[Z_i Z_j] = 0$ for $i \neq j$ and $EZ_i^2 = L2^{-2n}$

for some constant L. Using the monotone convergence theorem, we get

$$\begin{aligned}E\sum_{n\geq 1}(\sum_{t_i^n\leq t} Z_i)^2 &= \lim_{N\to\infty}\sum_{n\leq N} E(\sum_{t_i^n\leq t} Z_i)^2 \\ &= \lim_{N\to\infty}\sum_{n\leq N}(\sum_{t_i^n\leq t} L2^{-2n}) = L\sum_{n\geq 1} t2^{-n} < +\infty.\end{aligned}$$

Therefore $\sum_{n\geq 1}(\sum_{t_i^n\leq t} Z_i)^2$ is almost surely finite, and thus $(\sum_{t_i^n\leq t} Z_i)^2$ converge to zero (if an infinite sum of real numbers convergences, then the summands have to converge to zero). But then also $\sum_{t_i^n\leq t} Z_i$ converges to zero, so we have shown that $QV_t^n(W)\to t$ almost surely.

Since D is countable, we can find a P-null-set N such that $QV_s^n(W)\to s$ for every $s\in D$ and $\omega\notin N$. Since, for each n, $t\mapsto QV_t^n(W)$ is increasing, we get, for $s^\nu\in D$ increasing to t, $t^\nu\in D$ decreasing to t, and for every $\omega\notin N$, that

$$t=\lim_\nu s^\nu=\lim_\nu\lim_n QV_{s^\nu}^n(W)\leq\lim QV_t^n(W)\leq\lim_\nu\lim_n QV_{t^\nu}^n(W)=\lim_\nu t^\nu=t$$

$\square$

Lemma 2.15. *Assume that z is an adapted continuous stochastic process with $\sup_t Ez_t^2<\infty$. Then, for every t,*

$$\lim\sum_{t_i^{(n)}\leq t}(z_{t_i^{(n)}}(W_{t_{i+1}^{(n)}}-W_{t_{i+1}^{(n)}})^2)=\int_0^t z_s ds,$$

where the convergence is in L^2.

Proof. We denote $\eta_i:=z_{t_i^{(n)}}$, $L=\sup_t Ez_t^2<\infty$, $\Delta t_{i+1}^{(n)}:=t_{i+1}^{(n)}-t_i^{(n)}$ and $\Delta W_{i+1}^{(n)}:=W^{(n)}_{t_{i+1}^{(n)}}-W^{(n)}_{t_i^{(n)}}$ so that

$$\sum_{t_i^{(n)}\leq t}\left(z_{t_i^{(n)}}(W_{t_{i+1}^{(n)}}-W_{t_i^{(n)}})^2-z_{t_i^{(n)}}(t_{i+1}^{(n)}-t_i^{(n)})\right)=\sum_{t_i^{(n)}\leq t}\eta_i((\Delta W_{i+1}^{(n)})^2-\Delta t_{i+1}^{(n)}).$$

Recalling that the increments of Brownian motion are independent of the past and z is adapted, we get from $E[(\Delta W_{i+1}^{(n)})^2-\Delta t_{i+1}^{(n)})]=0$ and independence that

$$\begin{aligned}&E\left[\eta_i(\Delta W_{i+1}^{(n)})^2-\Delta t_{i+1}^{(n)})\eta_j(\Delta W_{j+1}^{(n)})^2-\Delta t_{j+1}^{(n)})\right]\\ =&E\left[(\Delta W_{i+1}^{(n)})^2-\Delta t_{i+1}^{(n)})\right]\left[E\eta_i\eta_j(\Delta W_{j+1}^{(n)})^2-\Delta t_{j+1}^{(n)})\right]\\ =&0\end{aligned}$$

for $i > j$. Combining with Remark 2.9, we get for some constants L (differing from line to line),

$$
\begin{aligned}
E|\sum_{t_i^{(n)}\leq t} \eta_i((\Delta W_{i+1}^{(n)})^2 - \Delta t_{i+1}^{(n)})|^2 &= \sum_{t_i^{(n)}\leq t} E|\eta_i((\Delta W_{i+1}^{(n)})^2 - \Delta t_{i+1}^{(n)})|^2 \\
&\leq L \sum_{t_i^{(n)}\leq t} E((\Delta W_{i+1}^{(n)})^2 - \Delta t_{i+1}^{(n)})^2 \\
&\leq L \sum_{t_i^{(n)}\leq t} \left(\Delta t_{i+1}^{(n)}\right)^2 \\
&= Lt2^n(2^{-2n}) \\
&\to 0.
\end{aligned}
$$

Since $\sum_{t_i^{(n)}\leq t} z_{t_i^{(n)}} \Delta t_{i+1}^n$ converges to $\int_0^t z_s ds$ in L^2, the claim follows from the triangle inequality. □

Remark 2.16. *Choosing $z = 1$ in Lemma 2.15, we get $QV^n(W)_t \to t$ in L^2.*

2.6 Stochastic integrals

Theorem 2.14 implies that the paths of the Brownian motion are not of bounded variation, and thus not differentiable. Indeed,

$$QV_t^n(W) \leq \max_{i\geq 1} |W_{t_{i+1}^{(n)}} - W_{t(n)_i}| \sum_{i\geq 1} |W_{t(n)_{i+1}} - W_{t(n)_i}|,$$

where, almost surely, $QV_t^n(W)$ converge to t and $\max_{i\geq 1} |W_{t(n)_{i+1}} - W_{t(n)_i}|$ converges to zero (by continuity of BM), so $\sum_{i\geq 1} |W_{t(n)_{i+1}} - W_{t(n)_i}|$ has to converge to $+\infty$. This means that it is not possible integrate functions with respect to the paths of Brownian motion in the usual sense of the Lebesque-Stieltjes integration theory.

However, it is possible to define integrals with respect to the Brownian motion in the sense of stochastic integrals. An adapted stochastic process z is *simple* if

$$z_t = \sum_{i=0}^{\infty} \eta_i 1_{(t_i,t_{i+1}]}(t)$$

for some $0 \leq t_1 \leq t_2 \leq \ldots$ and $\mathcal{F}_{t_i}$-measurable η_i with $\sup_i \operatorname{ess\,sup} |\eta_i| < \infty$. For a simple z, we set

$$\int_0^t z_t dW_t := \sum_{i=0}^{T} \left(\eta_i(W_{t\wedge t_{i+1}} - W_{t\wedge t_i})\right).$$

We extend the definition from simple processes to larger spaces of integrands

$$\mathcal{H}^2 := \{z \mid z \text{ measurable adapted stochastic process}, E\int_0^T |z_t|^2 dt < \infty\},$$

which we equip with the norm $\|z\|_{\mathcal{H}^2} := (E\int_0^T |z_t|^2 dt)^{1/2}$. For $z \in \mathcal{H}^2$, we define the stochastic integral as the unique limit

$$\int_0^t z_s dW_s := \lim_n \int_0^t z_s^{(n)} dW_s$$

in L^2, where $(z^{(n)})$ is any sequence of simple processes converging to z in $\mathcal{H}^2$.

Example 2.17. *Let $z = 2W$. For $\{t_0^n, t_i^n, \dots\} = D^n$, it is possible to show that the processes $z_t^{(n)} = \sum_{i\geq 0} 2W_{t_i^n} 1_{(t_i^n, t_{i+1}^n]}(t)$ converge to W in $\mathcal{H}^2$. We have*

$$\begin{aligned}
\int_0^t z_s^{(n)} dW_s &:= \sum_{i\geq 0} \Big(2W_{t_i^n}(W_{t\wedge t_{i+1}^n} - W_{t\wedge t_i^n}) \Big) \\
&= \sum_{i\geq 1} (W^2_{t\wedge t_{i+1}^n} - W^2_{t\wedge t_i^n}) - \sum_{i\geq 1} (W_{t\wedge t_{i+1}^n} - W_{t\wedge t_i^n})^2 \\
&= W_t^2 - QV_t^{(n)}(W) \\
&\to W_t^2 - t,
\end{aligned}$$

where the convergence is in L^2, by Remark 2.16. Thus

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

Theorem 2.18. *Let $z \in \mathcal{H}^2$ and S be the stochastic process defined by*

$$S_t = \int_0^t z_s dW_s.$$

1. *The process S is a continuous martingale that belongs to $\mathcal{H}^2$,*
2. *We have the* Itô isometry $ES_T^2 = \|z\|^2_{\mathcal{H}^2}$,
3. *If z is deterministic (and $\int_0^T |z_s|^2 ds < \infty$), then S has independent increments and $(S_t - S_s) \sim N(0, \int_s^t |z_u|^2 du)$.*

Next we extend the definition of the stochastic integral to integrands in the space $\mathcal{H}^2_{\text{loc}}$, where

$$\mathcal{H}^p_{\text{loc}} := \{z \mid z \text{ measurable adapted stochastic process}, \int_0^T |z_t|^p dt < \infty \ P\text{-a.s.}\}.$$

For $z \in \mathcal{H}^2_{\text{loc}}$, we define the stochastic integral as the unique limit

$$\int_0^t z_s dW_s := \lim_n \int_0^t z_s^{(n)} dW_s$$

where $z_t^{(n)} = z_t \mathbb{1}_{t \wedge \tau^{(n)}}$ and $\tau^{(n)} = \inf_t \{\int_0^t |z_t|^2 dt \geq n\}$ (here $\tau^{(n)}$ is a "localizing sequence of z). The stochastic process defined via $\int_0^t z_t dW_t$ is a continuous process, but not necessarily a martingale (it is only a "local martingale").

2.7 Exercises

Exercise 2.7.1. *Show that a Brownian motion W is a martingale with respect to its natural filtration $\mathcal{F}_t = \sigma(W_s \mid s \leq t)$, that is, for all $s < t \leq T$, $s \geq 0$, we have*

$$E[W_t \mid \mathcal{F}_s] = W_s.$$

Exercise 2.7.2. *Using Remark 2.9, show that, for any $\epsilon > 0$, Brownian motion has $(1/2 - \epsilon)$-Hölder continuous paths almost surely.*

Exercise 2.7.3. *Recall the continuous mapping theorem: If $(\eta_0^\nu, \ldots, \eta_k^\nu) \xrightarrow{d} (\eta_0, \ldots, \eta_k)$, then $f(\eta_0^\nu, \ldots, \eta_k^\nu) \xrightarrow{d} f(\eta_0, \ldots, \eta_k)$ for any continuous function $f : \mathbb{R}^k \to \mathbb{R}^n$.*

Use the continuous mapping theorem to finish the proof of Theorem 2.13.

Exercise 2.7.4. *In the setting of Example 2.17, show that $z^{(n)} \to z$ in $\mathcal{H}^2$.*

Exercise 2.7.5. *Let S be a stochastic process defined by*

$$S_t = \exp\{\int_0^t z_s dW_s - \frac{1}{2} \int_0^t |z_s|^2 ds\},$$

where z is deterministic with $\int_0^T |z_s|^2 ds < \infty$.

- *Compute $E \exp(\eta)$ for $\eta \sim N(0, \sigma^2)$, where $\sigma \in \mathbb{R} > 0$.*
- *Show that S is a martingale without relying on the first part of Theorem 2.18.*

2.8 Itô processes and Itô's formula

An important difference to the classical integration theory is that the stochastic integral does not satisfy the usual chain rule. Recall that for continuously differentiable functions g on $\mathbb{R}$ and f on $[0, T]$, we have $\frac{d}{dt} g(f) = g'(f) f'$ and so

$$g(f_t) = g(f_0) + \int_0^t g'(f_s) df_s.$$

Example 2.17 shows that this is not the case for the stochastic integral, since we got

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

where we have an "Itô correction term" involving the quadratic variation of W. This observation generalizes to the famous Ito's formula that we formulate directly to Itô processes.

Definition 2.19. *A stochastic process X is called an* Itô process*, if there is $\mu \in \mathcal{H}^1_{loc}$ and $\sigma \in \mathcal{H}^2_{loc}$ such that*

$$X_t = X_0 + \int_0^t \mu_s ds + \int \sigma_s dW_s.$$

The definition of the stochastic integral extends to Itô processes. Let

$$X_t = X_0 + \int_0^t \mu_s ds + \int \sigma_s dW_s$$

be an Itô process with $\mu \in \mathcal{H}^1_{\mathrm{loc}}$ and $\sigma \in \mathcal{H}^2_{\mathrm{loc}}$. For any z such that $z\mu \in \mathcal{H}^1_{\mathrm{loc}}$ and $z\sigma \in \mathcal{H}^2_{\mathrm{loc}}$, we define

$$\int_0^t z_s dX_s := \int_0^t z_s \mu_s ds + \int_0^t z_s \sigma_s dW_s.$$

We denote by $C^{1,2}$ the continuous functions $(t,x) \mapsto f(t,x)$ on $[0,T] \times \mathbb{R}$ that are continuously differentiable once w.r.t t and twice w.r.t. x.

Theorem 2.20 (Itô's formula)**.** *Assume that $f \in C^{1,2}([0,T] \times \mathbb{R})$ and that*

$$X_t = X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s$$

for some $\mu \in \mathcal{H}^1_{loc}$ and $\sigma \in \mathcal{H}^2_{loc}$. Then, almost surely,

$$f(t, X_t) = f(0, X_0) + \int_0^t \partial_x f(s, X_s) dX_s + \int_0^t (\partial_t f(s, X_s) + \frac{1}{2}\sigma_s^2 \partial_{xx} f(s, X_s)) ds.$$

Proof. We do not give the whole proof, but we only demonstrate how the "correction term" $\frac{1}{2}\sigma_s^2 \partial_{xx} f(s, X_s))ds$ appears to the formula in the special case when $X = W$, f is constant w.r.t. t-component, $f(0) = 0$, and $\partial_x f$ and $\partial_{xx} f$ are bounded. For general X, the argument follows similarly while boundedness of the derivatives can be handled using localizing sequences of X. Using Taylor's expansion (below $\eta_i^{(n)}$ is the appropriate random variable with

$W_{t_{i+1}^{(n)}} \leq \eta_i^{(n)} \leq W_{t_{i+1}^{(n)}}$), we get

$$
\begin{aligned}
(f(W_{t_{i+1}^{(n)}} - f(W_{t_i^{(n)}})) = \partial_x f(W_{t_i^{(n)}})(W_{t_{i+1}^{(n)}} - (W_{t_i^{(n)}}) \\
+ \frac{1}{2}\partial_{xx} f(W_{t_i^{(n)}})(W_{t_{i+1}^{(n)}} - W_{t_i^{(n)}})^2 \\
+ \frac{1}{2}(\partial_{xx} f(\eta_i^{(n)}) - \partial_{xx} f(W_{t_i^{(n)}}))(W_{t_{i+1}^{(n)}} - W_{t_i^{(n)}})^2.
\end{aligned}
$$

Summing over i we arrive at

$$
\begin{aligned}
f(W_t) = &\sum_n \partial_x f(W_{t_i^{(n)}})(W_{t_{i+1}^{(n)}} - W_{t_i^{(n)}}) \\
&+ \sum_n \frac{1}{2}\partial_{xx} f(W_{t_i^{(n)}})(W_{t_{i+1}^{(n)}} - W_{t_i^{(n)}})^2 \\
&+ \sum_n \frac{1}{2}(\partial_{xx} f(\eta_i^{(n)}) - \partial_{xx} f(W_{t_i^{(n)}}))(W_{t_{i+1}^{(n)}} - W_{t_i^{(n)}})^2.
\end{aligned}
$$

As n tends to infinity, the first sum converges to $\int_0^t \partial_x f(W_s)dW_s$, the second sum converges to $\int_0^t \partial_{xx} f(W_s)ds$, by Lemma 2.15, and the last term tends to zero (for this we omit the details). □

We will also need the following (local) martingale representation result.

Theorem 2.21 (The martingale representation theorem)**.** *Assume that $\eta \in L^1(\mathcal{F}_T)$ and that the filtration is generated by a Brownian W. Then there exists $z \in \mathcal{H}^2_{loc}$ such that, almost surely,*

$$
\eta = E[\eta] + \int_0^T z_t dW_t.
$$

If $\eta \in L^2(\mathcal{F}_T)$, then $z \in \mathcal{H}^2$ and z in the above representation is unique.

2.9 Multi-dimensional results

We denote the $\mathbb{R}^{d\times d}$-identity matrix by 1_d.

Definition 2.22. *An $\mathbb{R}^d$-valued stochastic process W is a d-*dimensional standard Brownian motion *if the components W^i are independent standard Brownian motions and $W_t - W_s \sim N(0, (t-s)1_d)$.*

Theorem 2.23. *Let W be a d-dimensional standard Brownian motion. Then*

$$
\sum_{t_i^n \leq t} (W_{t_{i+1}^n} - W_{t_t^n})(W_{t_{i+1}^n} - W_{t_t^n})^T \xrightarrow{n\to\infty} t1_d
$$

where the convergence is in L^2.

Definition 2.24. *A stochastic process X is called a d-*dimensional Itô process*, if there is $\mu \in \mathcal{H}^1_{loc}(\mathbb{R}^d)$ and $\sigma \in \mathcal{H}^2_{loc}(\mathbb{R}^d \times \mathbb{R}^n)$ such that*

$$X^i_t = X^i_0 + \int_0^t \mu^i_s ds + \int \sigma^i_s dW_s,$$

where σ^i_s is the i-th row of the matrix σ_s and W is an n-dimensional standard Brownian motion.

Theorem 2.25 (Multi-dimensional Itô's formula)**.** *Assume that $f \in C^{1,2}([0,T] \times \mathbb{R}^d)$ and that*

$$X_t = X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_t dW_s$$

for some a n-dimensional standard Brownian motion and $\mu \in \mathcal{H}^1_{loc}(\mathbb{R}^d)$ and $\sigma \in \mathcal{H}^2_{loc}(\mathbb{R}^{d\times n})$. Then, almost surely,

$$f(t, X_t) = f(0, X_0) + \int_0^t \partial_x f(s, X_s) dX_s + \int_0^t (\partial_t f(s, X_s) + \frac{1}{2} \operatorname{Tr}[\partial_{xx} f(s, X_s)\sigma_s \sigma_s^T]) ds.$$

Here

$$\int_0^t \partial_x f(s, X_s) dX_s := \sum_{i=1}^d \int_0^t \partial_{x^i} f(s, X_s) dX^i_s.$$

Exercise 2.9.1 (Integration by parts formula for Itô processes)**.** *Let W be one dimensional Brownian motion and let X^i, $i = 1, 2$ be one dimensional Itô processes, i.e.,*

$$X^i_t = X^i_0 + \int_0^t \mu^i_s ds + \int_0^t \sigma^i_s dW_s \quad i = 1, 2,$$

where, for $i = 1, 2$, $\mu^i \in \mathcal{H}^1_{loc}$ and $\sigma^i \in \mathcal{H}^2_{loc}$. Apply Itô's formula to $f(t, (x^1, x^2)) = x^1 x^2$ to show the integration by parts formula

$$X^1_t X^2_t = X^1_0 X^2_0 + \int_0^t X^2_s dX^1_s + \int_0^t X^1_s dX^2_s + \int_0^t \sigma^1_s \sigma^2_s ds.$$

Exercise 2.9.2. *Let W be a d-dimensional Brownian motion, $b \in \mathbb{R}$ and $\sigma \in \mathbb{R}^d$, and consider the process*

$$S_t = \exp\{bt + \sigma \cdot W_t\}.$$

For $f \in C^{1,2}([0,T] \times \mathbb{R})$, show that $f(t, S_t)$ is an Itô process. Can you find a function f such that the process given by $Y_t := f(t, S_t)$ is a local martingale?

2.10 Exercises

We say that the process defined by

$$B_t := \mu t + \sigma W_t$$

is a Brownian motion with drift $\mu \in \mathbb{R}^d$ and volatility $\sigma \in \mathbb{R}^{d\times d}$, where W is a (standard) d-dimensional Brownian motion W. The matrix $Cov = \sigma\sigma^T$ is called the *covariation matrix*.

Instead of specifying σ, it is more common to specify Cov which can be estimated from data. It is an exercise to check that volatility matrices with a common covariation matrix define Brownian motions with common finite dimensional distributions.

Example 2.26. *Assume that $\eta = (\eta_1, \ldots, \eta_d)$ is a random vector of independent standard distributions and $\Sigma \in \mathbb{R}^{d\times d}$ be given. Let $\sigma \in \mathbb{R}^{d\times d}$ be such that $\sigma\sigma^T = \Sigma$. Then*

$$\mathrm{Cov}(\sigma\eta) = E[(\sigma\xi)(\sigma\eta)^T] = \sigma E[\eta\eta T]\sigma^T = \sigma\sigma^T = \Sigma.$$

Thus the covariation depends on σ only through $\sigma\sigma^T$. We can always use the Cholesky decomposition $\Sigma = LL^T$, where L is a triangular matrix.

For B_t this implies the following. When only the covariation is specified, one can always use L from the Cholesky decomposition as σ. Given any other $\hat{\sigma}$ that gives the same covariation, the processes B and

$$\hat{B}_t := \mu t + \hat{\sigma} W_t$$

have the same finite dimensional disributions.

Exercise 2.10.1. *Write a Matlab function (as an .m-file) that creates a sample path of a d-dimensional Brownian motion with terminal time T, $n+1$ equi-distant discretization points, drift vector μ and covariation matrix Cov. Write it as a function of these parameters and an i.i.d. sample of d-dimensional standard normals. Hint: Use Cholesky decomposition of Cov.*

Exercise 2.10.2. *Plot a sample path of a 2-dimensional Brownian motion with terminal time T, $n+1$ equi-distant discretization points with a drift vector μ and covariation matrix Cov. Plot it as a 3D graph, a 2D parametric curve (time being the parameter), and as each component as a different curve in the same figure. Which plot is the most informative?*

Exercise 2.10.3. *Plot sample paths of a 3-dimensional Brownian motion with terminal time T, $n+1$ equi-distant discretization points with a drift vector μ and with different covariation matrices Cov. Plot it so that each component is a different curve in the same figure. Vary Cov so that the role of covariation matrix becomes clear in the figures.*

Exercise 2.10.4. *For the path in "bmpath3D.mat" of a 3D Brownian motion, estimate the covariation matrix. Plot the components of the path in the same figure and label the paths with the rows of the covariation matrix (with the precision of two digits).*