Chapter 1: Crash course in probability theory

- A triple $(\Omega, \mathscr{A}, \mathbb{P})$ is called a probability space (PS) provided that:
  - the sample space $\Omega \neq \emptyset$ (this is the set of all possible outcomes);
  - the set of events $\mathscr{A}$ is a σ-algebra of subsets of Ω with

    (σ_1) $\Omega \in \mathscr{A}$

    (σ_2) If $A \in \mathscr{A}$ then its complement $A^{\mathsf{C}} \in \mathscr{A}$, where $A^{\mathsf{C}} = \Omega \setminus A$

    (σ_3) If $A_1, A_2, A_3, \ldots \in \mathscr{A}$ then $\bigcup_{n=1}^{\infty} A_n \in \mathscr{A}$;
  - the map $\mathbb{P}\colon \mathscr{A} \to [0,1]$ is a probability measure, that is

    (p_1) $\mathbb{P}(\Omega) = 1$

    (p_2) $\mathbb{P}(A^{\mathsf{C}}) = 1 - \mathbb{P}(A)$ for all $A \in \mathscr{A}$

    (p_3) If $A_1, A_2, A_3, \ldots \in \mathscr{A}$ are disjoint (i. e. $A_i \cap A_j = \emptyset$ for $i \neq j$), then
    $\mathbb{P}(\bigcup_{n=1}^{\infty} A_n) =: \mathbb{P}(\biguplus_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$.

- Two events $A, B \in \mathscr{A}$ are independent if $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$.

- A property which is true except for an event of probability zero is said to hold almost surely (a.s).

- The σ-algebra of Borel sets in $\mathbb{R}$ is denoted by $\mathscr{B}(\mathbb{R})$. It is the smallest σ-algebra containing all intervals of the form $[a,b[$ for reals $a<b$ ($]-\infty,b[$ is also ok). In a similar way, $\mathscr{B}([0,1])$ denotes the smallest σ-algebra containing all intervals of the form $[a,b[$, for reals $a<b$ with $[a,b[\subset [0,1]$.

  Observe that the Borel σ-algebra does not contain only intervals of the form $[a,b[$, but also (for example) intervals of the form $[a,b]$, $]a,b[$, $]a,b]$, or the singleton $\{a\}$.

- Consider $(\Omega, \mathscr{A}, \mathbb{P})$ a PS. A (real-valued) random variable (RV) is a measurable function $X\colon \Omega \to \mathbb{R}$. Here, measurable means $X^{-1}(B) \in \mathscr{A}$ for all $B \in \mathscr{B}(\mathbb{R})$, where $X^{-1}(B) = \{\omega \in \Omega\colon\ X(\omega) \in B\}$.

- A RV X on a PS $(\Omega, \mathscr{A}, \mathbb{P})$ is called continuous if there exists a piecewise continuous non-negative function $p_X\colon \mathbb{R} \to [0,1]$ such that the cumulative distribution function of X (CDF), $F_X\colon \mathbb{R} \to [0,1]$, satisfies
  $$F_X(x) := \mathbb{P}(X \leq x) := \mathbb{P}(\{\omega \in \Omega\colon\ X(\omega) \leq x\}) = \int_{-\infty}^{x} p_X(s)\,\mathrm{d}s \quad \text{for all} \quad x \in \mathbb{R}.$$
  In this case, we call p_X the probability density function of X (PDF).

  From this definition, it follows that $\mathbb{P}(X \in B) = \int_B p_X(s)\,\mathrm{d}s$ for all Borel sets $B \in \mathscr{B}(\mathbb{R})$. Furthermore, $\mathbb{P}(a \leq X \leq b) = F_X(b) - F_X(a) = \int_a^b p_X(s)\,\mathrm{d}s$ for reals $a<b$. Finally, if p_X is continuous at $x \in \mathbb{R}$, one has $p_X(x) = \frac{\mathrm{d}}{\mathrm{d}x} F_X(x)$.

- Let $-\infty < a < b < \infty$. A RV X is uniformly distributed in $[a,b]$ if its probability density function is given by
  $$p_X(x) = \begin{cases} \frac{1}{b-a} & \text{for} \quad x \in [a,b] \\ 0 & \text{else.} \end{cases}$$
  Notation: $X \sim \mathscr{U}(a,b)$.

- Let $\mu, \sigma \in \mathbb{R}$ with $\sigma^2 > 0$. A RV X is normally distributed or a Gaussian random variable if its probability density function is given by

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

  If $\mu = 0$ and $\sigma = 1$, then X has a standard normal distribution or is a standard Gaussian random variable.

  Notation: $X \sim \mathcal{N}(\mu, \sigma^2)$.

- Let X be a continuous RV with probability density function p. Let Y be another continuous RV. Let $g\colon \mathbb{R} \to \mathbb{R}$ and let a positive integer k.

  The expected value of X or mean of X or expectation of X is defined as

$$\mu := \mathbb{E}[X] := \int_\Omega X(\omega)\,\mathrm{d}\mathbb{P}(\omega) = \int_{-\infty}^{\infty} x\,p(x)\,\mathrm{d}x.$$

  Similarly, we define

$$\mathbb{E}[g(X)] := \int_{-\infty}^{\infty} g(x)p(x)\,\mathrm{d}x$$

  and the kth moment of X

$$\mathbb{E}[X^k] := \int_{-\infty}^{\infty} x^k p(x)\,\mathrm{d}x$$

  as well as the variance of X

$$\mathrm{Var}[X] = \mathbb{E}[(X-\mu)^2] = \mathbb{E}[X^2] - \mu^2.$$

  Finally, the covariance of X and Y is given by

$$\mathrm{Cov}(X,Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])].$$

  If $\mathrm{Cov}(X,Y) = 0$, one says that the RV X and Y are uncorrelated.

  Examples: For $X \sim \mathcal{U}(a,b)$, one has $\mathbb{E}[X] = \frac{1}{2}(a+b)$ and $\mathrm{Var}[X] = \frac{1}{12}(b-a)^2$. For $X \sim \mathcal{N}(\mu,\sigma^2)$, one has $\mathbb{E}[X] = \mu$ and $\mathrm{Var}[X] = \sigma^2$.

- Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a PS. For $A \in \mathcal{A}$, we define the indicator function or characteristic function $I_A = \chi_A\colon \Omega \to \mathbb{R}$ by

$$I_A(\omega) = \chi_A(\omega) = \begin{cases} 1 & \text{if} \quad \omega \in A \\ 0 & \text{else.} \end{cases}$$

  Observe that this is a RV. A simple random variable X has the form

$$X(\omega) = \sum_{j=1}^n c_j I_{A_j}(\omega),$$

  where $n \in \mathbb{N}, c_j \in \mathbb{R}, A_j \in \mathcal{A}$.

- The space of simple random variables is denoted by S_{RV}. It is equipped with the inner product $(X,Y) := \mathbb{E}[XY]$ and norm $\|X\|_{RV} := (X,X)^{1/2} = (\mathbb{E}[X^2])^{1/2}$ for $X, Y \in S_{RV}$. The space S_{RV} is dense in its completion the Hilbert space of random variables H_{RV}.

- Consider a sequence of RV $\{X_n\}_{n=1}^\infty$ and a RV X defined on a PS $(\Omega, \mathcal{A}, \mathbb{P})$. Let $p > 0$. We have the following types of convergence:

- $\{X_n\}_{n=1}^\infty$ converge strongly in L^2 or in the mean-square sense to X if

$$\mathbb{E}[|X_n - X|^2] = \|X_n - X\|_{RV}^2 \longrightarrow 0 \quad \text{as} \quad n \longrightarrow \infty.$$

- $\{X_n\}_{n=1}^\infty$ converge strongly in L^1 or strongly to X if

$$\mathbb{E}[|X_n - X|] \longrightarrow 0 \quad \text{as} \quad n \longrightarrow \infty.$$

- $\{X_n\}_{n=1}^\infty$ converge strongly in L^p to X if

$$\mathbb{E}[|X_n - X|^p] \longrightarrow 0 \quad \text{as} \quad n \longrightarrow \infty.$$

- $\{X_n\}_{n=1}^\infty$ converge in probability to X if

$$\forall \varepsilon > 0, \quad \lim_{n\to\infty} \mathbb{P}(|X_n - X| \geq \varepsilon) = 0.$$

- $\{X_n\}_{n=1}^\infty$ converge to X almost surely (a.s.) or with probability 1 if

$$\mathbb{P}(\{\omega \in \Omega\colon \lim_{n\to\infty} |X_n(\omega) - X(\omega)| = 0\}) = 1.$$

- $\{X_n\}_{n=1}^\infty$ converge in distribution to X if

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x) \quad \text{at all points, where } F_X \text{ is continuous.}$$

- $\{X_n\}_{n=1}^\infty$ converge weakly to X if

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} f(x) F_{X_n}(x)\,\mathrm{d}x = \int_{-\infty}^{\infty} f(x) F_X(x)\,\mathrm{d}x \quad \text{for all smooth functions } f.$$

In general one has: a.s. convergence $\Rightarrow$ convergence in probability $\Rightarrow$ convergence in distribution $\Leftrightarrow$ weak convergence. And ms convergence $\Rightarrow$ convergence in probability.

- For $0 < p < r$, Lyapunov inequality reads

$$(\mathbb{E}[|X|^p])^{1/p} \leq (\mathbb{E}[|X|^r])^{1/r}.$$

- For $\varphi\colon \mathbb{R} \to \mathbb{R}$ convex, Jensen's inequality reads

$$\varphi(\mathbb{E}[X]) \leq \mathbb{E}[\varphi(X)].$$

- If $\sum\limits_{n=1}^{\infty} \mathbb{P}(|X - X_n| \leq \varepsilon) < \infty$ for all $\varepsilon > 0$, then $X_n \longrightarrow X$ a.s.

Further resources:

- https://www.cs.utah.edu/~fletcher/cs6957/lectures/ProbabilityCrashCourse.pdf
- https://faculty.math.illinois.edu/~kkirkpat/SampleSpace.pdf
- https://onlinecourses.science.psu.edu/stat414/node/5
- http://www.randomservices.org/random/prob/index.html

- https://www.probabilitycourse.com/
- https://www.statlect.com/fundamentals-of-probability/

Chapter 2: Stochastic processes

This chapter introduces stochastic processes. This is needed to define, for instance, solutions to SDEs. Below, we let $(\Omega, \mathscr{A}, \mathbb{P})$ be a fixed probability space.

- A stochastic process (SP) is a family of random variables $\{X(t), t \in \tau\} = \{X(t)\}_{t\in\tau}$ defined on the probability space $(\Omega, \mathscr{A}, \mathbb{P})$ and indexed by a parameter t varying over a set τ.

  Observe, that $X\colon \tau \times \Omega \to \mathbb{R}$, so that (notation) $X := X(t) := X(t,\omega)$. Remark that $X(t,\cdot)$ is a RV for each fixed $t \in \tau$. And, for a fixed $\omega \in \Omega$, $X(\cdot,\omega)\colon \tau \to \mathbb{R}$ is called the sample path, realisation, trajectory of the SP $X(t)$.

  If τ is discrete, e.g. $\tau = \{0,1,2,3\}$ or $\tau = \mathbb{N}$, then the SP is called discrete. The discrete random walk is an example of a (discrete) stochastic process.

  If τ is continuous, e.g. $\tau = [0,1]$ or $\tau = \mathbb{R}$, then the SP is called continuous. The Brownian motion (BM) also called Wiener process (WP) is an example of a (continuous) stochastic process (see below).

- Let $T > 0$. A standard one-dimensional Brownian motion (BM), also called Wiener process (WP), on $[0,T]$ is a real-valued, with a.s. continuous sample paths, SP $\{B(t)\}_{t\in[0,T]}$ such that

(BM1) $B(0) = 0$ almost surely;

(BM2) For all $s < t$ with $s,t \in [0,T]$, the increment $B(t) - B(s)$ is normally distributed with mean zero and variance $t-s$, i. e. $B(t) - B(s) \sim \mathcal{N}(0, t-s)$;

(BM3) $B(t)$ has independent increments: for all $0 \le t_1 < t_2 \le t_3 < t_4 \le T$, $B(t_2) - B(t_1)$ and $B(t_4) - B(t_3)$ are independent.

  The notation $W(t)$ is also used for a BM/WP.

  A BM $\{B(t)\}_{t\in[0,T]}$ has the following properties:

  - $\mathbb{E}[B(t)] = 0$ for all $t \in [0,T]$.
  - $\mathrm{Cov}(B(t), B(s)) = s \wedge t = \min(s,t)$ for all $s,t \in [0,T]$.
  - $B(t)$ is almost surely nowhere differentiable.

- An elementary stochastic process, also called a random step function, $f := f(t,\omega) := \{f(t)\}_{t\in[0,T]}$ has the form

$$f(t,\omega) = \sum_{n=0}^{N-1} f(t_n,\omega) I_{[t_n,t_{n+1}[}(t),$$

  where $N \in \mathbb{N}$ is fixed, $0 = t_0 < t_1 < t_2 < \ldots < t_N = T$, $f(t_n,\cdot)$ are RV in H_{RV} for each fixed t_n, and $I_{[t_n,t_{n+1}[}$ denotes the (deterministic) characteristic function on the interval $[t_n, t_{n+1}[$.

  The linear space of all elementary SP (with $\displaystyle\int_0^T \mathbb{E}[f^2(t)]\,\mathrm{d}t = \sum_{n=0}^{N-1} \mathbb{E}[f^2(t_n,\omega)](t_{n+1} - t_n) < \infty$) will be denoted by S_{SP}. It is equipped with the inner product $(f,g)_{SP} := \displaystyle\int_0^T \mathbb{E}[f(t)g(t)]\,\mathrm{d}t$ which defines a norm $\|f\|_{SP} = (f,f)_{SP}^{1/2} = \Big(\displaystyle\int_0^T \mathbb{E}[f^2(t)]\,\mathrm{d}t\Big)^{1/2}$.

The completion of the space S_{SP} is denoted by H_{SP} and is an Hilbert space. Observe that S_{SP} is dense in H_{SP}.

Further resources:

- https://www.probabilitycourse.com/chapter10/10_1_0_basic_concepts.php
- http://www.randomservices.org/random/prob/Processes.html
- https://www.kent.ac.uk/smsas/personal/lb209/files/notes1.pdf contains some nice examples
- http://www.maths.manchester.ac.uk/~gajjar/magicalbooks/risk/a_Intro_to_SDEs.pdf (first part for the moment)

Chapter 3: Stochastic integration

This chapter introduces stochastic integration. This is the final needed tool to define, for instance, SDEs. Below, we let $(\Omega, \mathscr{A}, \mathbb{P})$ be a fixed probability space, $a < b$ be two real numbers, and $a = t_0 < t_1 < t_2 < \ldots < t_{N-1} < t_N = b$ be a partition of the interval $[a, b]$ with $N \in \mathbb{N}$. Furthermore, $I_{[t_n,t_{n+1}[}$ denotes the characteristic/indicator function on the interval $[t_n, t_{n+1}[$. Finally, $W(t)$ will denote a Brownian motion/Wiener process defined on this probability space.

Below, we should also assume that the integrands (a stochastic process denoted by f here) satisfies

(c1) $f(a,\cdot) \in H_{RV}$

(c2) $\|f(t_1,\cdot) - f(t_2,\cdot)\|_{RV}^2 \leq K|t_1 - t_2|$ for all $t_1, t_2 \in [a, b]$ and a positive constant K

(c3) f is non-anticipating on $[a, b]$, meaning more or less, that $f(t,\omega)$ does not depend on information from time $\tilde{t}$ with $\tilde{t} > t$ (i. e. the integrand $f(t)$ must be independent of the later values $\{W(s)\}_{s>t}$ of the Brownian path).

- The integral of an elementary stochastic process $\{f(t,\omega)\}_{t\in[a,b]} \in S_{SP}$ of the form $f(t,\omega) = \sum_{n=0}^{N-1} f_n(\omega) I_{[t_n,t_{n+1}[}(t)$, where $f_n \in H_{RV}$, is defined as

$$J(f)(\omega) := \int_a^b f(t,\omega)\,\mathrm{d}t = \int_a^b f\,\mathrm{d}t = \sum_{n=0}^{N-1} f_n(\omega)(t_{n+1} - t_n).$$

  It is a random variable with $J(f) \in H_{RV}$ as $\|f\|_{RV} < \infty$ by the Cauchy-Schwarz inequality.

- The stochastic integral of a stochastic process $f \in H_{SP}$ is defined as

$$J(f)(\omega) := \int_a^b f(t,\omega)\,\mathrm{d}t = \int_a^b f\,\mathrm{d}t = \lim_{N\to\infty} \int_a^b f_N\,\mathrm{d}t = \lim_{N\to\infty} \sum_{n=0}^{N-1} f_n^{(N)}(\omega)(t_{n+1} - t_n),$$

  where $\{f_N\}_{N=1}^{\infty}$ is a Cauchy sequence in S_{SP} converging to f in the sense: $\|f_N - f\|_{SP} \to 0$ as $N \to \infty$.

- Ito's stochastic integral for an elementary process $f \in S_{SP}$ is defined as

$$I(f)(\omega) := \int_a^b f(t,\omega)\,\mathrm{d}W(t,\omega) := \int_a^b f\,\mathrm{d}W = \sum_{n=0}^{N-1} f_n(\omega)\Delta W_n,$$

  with the Wiener increment $\Delta W_n := W(t_{n+1},\omega) - W(t_n,\omega) \sim N(0, t_{n+1} - t_n)$.
  Observe that $\|I(f)\|_{RV}^2 = \|f\|_{SP}^2 < \infty$ and so $I(f) \in H_{RV}$.

- The Ito integral of a stochastic process $f \in H_{SP}$ is defined as

$$I(f)(\omega) := \int_a^b f(t,\omega)\,\mathrm{d}W(t,\omega) = \int_a^b f\,\mathrm{d}W = \lim_{N\to\infty}\int_a^b f_N\,\mathrm{d}W = \lim_{N\to\infty}\sum_{n=0}^{N-1} f_n^{(N)}(\omega)\left(W(t_{n+1}^{(N)}) - W(t_n^{(N)})\right),$$

where $\{f_N\}_{N=1}^{\infty}$ is a Cauchy sequence in S_{SP} converging to f in the sense: $\|f_N - f\|_{SP} \to 0$ as $N\to\infty$. The following properties hold for Ito's integral:

1. $I(cf+g) = cI(f)+I(g)$ for any $f,g \in H_{SP}$ and $c\in\mathbb{R}$.
2. $\mathbb{E}[I(f)] = 0$.
3. Ito's isometry: $\mathbb{E}\left[\left|\int_a^b f\,\mathrm{d}W\right|^2\right] = \int_a^b \mathbb{E}[|f|^2]\,\mathrm{d}t$.
4. $\mathbb{E}[I(f)I(g)] = \int_a^b \mathbb{E}[fg]\,\mathrm{d}t$.

Further resources:

- http://venus.unive.it/imef/themes/imef/images/files/doc0910/StochasticIntegral2008.pdf
- http://www.math.wsu.edu/math/faculty/lih/SDE-week6.pdf
- https://www.youtube.com/watch?v=1OnJ7t_4-nM

Chapter 4: Stochastic differential equations

In this chapter we combine all the tools seen so far in order to define and analyse stochastic differential equations (SDE). Below, we let $(\Omega,\mathscr{A},\mathbb{P})$ be a fixed probability space, $T>0$, $X_0\in H_{RV}$, $f,g\colon\mathbb{R}\times\mathbb{R}\to\mathbb{R}$ and $(W(t))_{t\in[0,T]}$ a Brownian motion/Wiener process defined on this probability space.

- An Ito stochastic differential equation on the interval $[0,T]$ has the form

$$X(t,\omega) = X_0(\omega) + \int_0^t f(s,X(s,\omega))\,\mathrm{d}s + \int_0^t g(s,\omega)\,\mathrm{d}W(s,\omega) \quad \text{for} \quad 0\le t\le T. \tag{SDE}$$

We will also use the (simpler) representation in differential form

$$\begin{aligned} \mathrm{d}X(t) &= f(t,X(t))\,\mathrm{d}t + g(t,X(t))\,\mathrm{d}W(t) \\ X(0) &= X_0. \end{aligned}$$

The function f is called the drift coefficient, the function g the diffusion coefficient.

- For $\mu,\sigma\in\mathbb{R}$, a SP X is called a geometric Brownian motion if it is the solution to the SDE

$$\begin{aligned} \mathrm{d}X &= \mu X\,\mathrm{d}t + \sigma X\,\mathrm{d}W \\ X(0) &= X_0. \end{aligned}$$

This equation models stock prices in the Black-Scholes model. Its exact solution reads $X(t) = X_0\exp\left((\mu-\frac{\sigma^2}{2})t + \sigma W(t)\right)$.

- For $b>0$ and $\sigma\in\mathbb{R}$, the Langevin equation is given by

$$\begin{aligned} \mathrm{d}X &= -bX\,\mathrm{d}t+\sigma\,\mathrm{d}W \\ X(0) &= X_0. \end{aligned}$$

  This models the motion of particles in a fluid.

  Its exact solution reads $X(t)=\mathrm{e}^{-bt}X_0+\sigma\int_0^t \mathrm{e}^{-b(t-s)}\,\mathrm{d}W(s)$.

- Let $\lambda,\mu,\sigma>0$. The mean-reverting square root process is the solution to the SDE

$$\begin{aligned} \mathrm{d}X &= \lambda(\mu-X)\,\mathrm{d}t+\sigma\sqrt{X}\,\mathrm{d}W \\ X(0) &= X_0. \end{aligned}$$

  This is an interest rate model, also known as the CIR model.

- Consider an integer $N>0$, define the stepsize $\Delta t:=\frac{T}{N}$ and consider a partition of the interval $[0,T]$ given by $0=t_0<t_1<\ldots<t_N=T$, where $t_n=n\Delta t$ for $n=0,1,\ldots,N$. We present some numerical methods for the approximation of solutions to (SDE). This provides numerical approximations $X_n\approx X(t_n)$ for $n=1,2,\ldots,N$.

  The Euler-Maruyama scheme (EM) is given by

$$\begin{aligned} X_{n+1} &= X_n+f(t_n,X_n)\Delta t+g(t_n,X_n)\Delta W_n \\ X_0 &= X(0), \end{aligned}$$

  where the Wiener increments $\Delta W_n=W(t_{n+1})-W(t_n)\sim N(0,\Delta t)$.

  The backward Euler-Maruyama scheme (BEM) is given by

$$\begin{aligned} X_{n+1} &= X_n+f(t_{n+1},X_{n+1})\Delta t+g(t_n,X_n)\Delta W_n \\ X_0 &= X(0). \end{aligned}$$

  The Milstein scheme (M) reads

$$\begin{aligned} X_{n+1} &= X_n+f(t_n,X_n)\Delta t+g(t_n,X_n)\Delta W_n+\frac{1}{2}g'(X_n)g(X_n)\left(\Delta W_n^2-\Delta t\right) \\ X_0 &= X(0). \end{aligned}$$

- For a function $\Phi\colon\mathbb{R}\to\mathbb{R}$, the weak error is

$$\mathrm{err}_{\Delta t}^{\mathrm{weak}}:=\sup_{0\leq t_n\leq T}\left|\mathbb{E}[\Phi(X_n)]-\mathbb{E}[\Phi(X(t_n))]\right|.$$

  A numerical method for (SDE) converges weakly if for any function Φ of a certain class

$$\mathrm{err}_{\Delta t}^{\mathrm{weak}}\to 0 \quad \text{as} \quad \Delta t\to 0.$$

  A numerical method for (SDE) has weak order of convergence p if there exist $K>0$ and $\Delta t^*>0$ such that

$$\mathrm{err}_{\Delta t}^{\mathrm{weak}}\leq K\Delta t^p$$

  for $0\leq\Delta t\leq\Delta t^*$.

  For the above mentioned numerical methods, one has (in general) $p_{EM}=1$, $p_{BEM}=1$, and $p_M=1$.

- The strong error is
$$\mathrm{err}_{\Delta t}^{\mathrm{strong}} := \sup_{0\leq t_n\leq T} \mathbb{E}[|X_n - X(t_n)|].$$
A numerical method for (SDE) converges strongly if
$$\mathrm{err}_{\Delta t}^{\mathrm{strong}} \to 0 \quad \text{as} \quad \Delta t \to 0.$$
A numerical method for (SDE) has strong order of convergence q if there exist $K > 0$ and $\Delta t^* > 0$ such that
$$\mathrm{err}_{\Delta t}^{\mathrm{strong}} \leq K\Delta t^q$$
for $0 \leq \Delta t \leq \Delta t^*$.
For the above mentioned numerical methods, one has (in general) $q_{EM} = 1/2$, $q_{BEM} = 1/2$, and $q_M = 1$.

- We have investigated the statistical error due to a Monte-Carlo algorithm for the approximation of $\mathbb{E}[X]$ of some random variable X:
  1. Compute M independent realisations of X, denoted by X^k for $k = 1, 2, \dots, M$.
  2. Compute the sample average $\mathbb{E}^M[X] := \frac{1}{M}\sum_{k=1}^{M} X^k$.

  We have found that $\mathbb{E}[X] - \mathbb{E}^M[X] = \mathcal{O}(\frac{1}{\sqrt{M}})$. In particular, when considering the weak error of EM scheme, one should take $M \approx N^2$ realisations in order to properly approximate the expectations.

- Let $f, g\colon [0, T]\times\Omega \to \mathbb{R}$ be nice processes, $T > 0$, and $X_0 \in H_{RV}$. A stochastic process X is called an Ito process if it satisfies
$$X(t) = X_0 + \int_0^t f(s)\,\mathrm{d}s + \int_0^t g(s)\,\mathrm{d}W(s) \quad \text{for} \quad 0 \leq t \leq T.$$
We also say, that X has the stochastic differential
$$\begin{aligned} \mathrm{d}X(t) &= f(t)\,\mathrm{d}t + g(t)\,\mathrm{d}W(t) \\ X(0) &= X_0. \end{aligned}$$
For such processes and a nice function $F\colon [0, T]\times\mathbb{R} \to \mathbb{R}$, Ito's formula tells us that
$$\mathrm{d}(F(t, X(t))) = \widetilde{f}(t, X(t))\,\mathrm{d}t + \widetilde{g}(t, X(t))\,\mathrm{d}W(t), \tag{IF}$$
where $\widetilde{f}(t,x) := \frac{\partial F}{\partial t}(t,x) + f(t)\frac{\partial F}{\partial x}(t,x) + \frac{1}{2}g(t)^2\frac{\partial^2 F}{\partial x^2}(t,x)$ and $\widetilde{g}(t,x) := g(t)\frac{\partial F}{\partial x}(t,x)$.
Equation (IF) has to be understood as an integral equation. Furthermore, formula (IF) is for instance useful to find exact solutions to particular SDEs or to compute moments of such solutions.

- We have seen that, if the coefficients of the stochastic differential equation (SDE) are non-anticipating, globally Lipschitz and satisfy a linear growth condition, then (SDE) has a unique (strong) global solution. Furthermore, the solution (a stochastic process) $X(t)$ is continuous (with respect to time t) and has bounded second moment.

- We also studied the convergence of Euler-Maruyama's scheme. In particular, under some technical assumptions, we saw that (in general) the strong order of convergence of this numerical method is (in general) $1/2$ and the weak order of convergence is 1 (for the class of test functions that are smooth and with at most polynomial growth).

Further resources:

- http://www.azimuthproject.org/azimuth/show/Stochastic+differential+equation
- https://www.stat.auckland.ac.nz/~geoff/talks/SDE-notes.pdf
- https://www.mimuw.edu.pl/~apalczew/CFP_lecture5.pdf
- http://people.math.sfu.ca/~tupper/days/m3.pdf
- https://ocw.mit.edu/courses/sloan-school-of-management/15-070j-advanced-stochastic-proces lecture-notes/MIT15_070JF13_Lec17.pdf