TMS150 & MSG400

STOCHASTIC DATA PROCESSING AND SIMULATION

SIMULATION OF STOCHATSIC PROCESSES WITH PYTHON, LAB 6

ANNIKA LANG

Throughout this project, let $(\Omega, \mathcal{A}, P)$ be a probability space.

1. Stochastic processes

This first section of the project is devoted to the introduction of stochastic processes in discrete and continuous time, their properties, and to first examples. Let us start with the definition of a stochastic process.

Definition 1.1. Let $\mathbb{T} \subset \mathbb{R}$ be a set. A *stochastic process* X with *parameter space* $\mathbb{T}$ is a measurable function $X : \Omega \times \mathbb{T} \to \mathbb{R}$.

In other words, a stochastic is a collection of random variables indexed by the set $\mathbb{T}$. We use $X := (X(t), t \in \mathbb{T})$ as equivalent notation.

The parameter t is called the *time parameter*. We distinguish between *discrete time*, where usually $\mathbb{T} = \mathbb{Z}$ or $\mathbb{T} = \mathbb{N}$, and *continuous time*, where usually $\mathbb{T} = \mathbb{R}$, $\mathbb{T} = \mathbb{R}_+$, or $\mathbb{T} = [0, T]$ for some $T < \infty$.

A plot of a stochastic process $X(\omega, t)$ as a function of $t \in \mathbb{T}$ for fixed $\omega \in \Omega$ is called a *sample path* of the process. When a stochastic process is plotted, it is a (hopefully typical) sample path that is depicted.

Definition 1.2. A stochastic process X has *expectation function* $m_X : \mathbb{T} \to \mathbb{R}$ given by

$$m_X(t) := \mathbb{E}[X(t)]$$

and *covariance function* $\gamma_X : \mathbb{T} \times \mathbb{T} \to \mathbb{R}$ given by

$$\gamma_X(s, t) := \mathsf{Cov}[X(s), X(t)] = \mathbb{E}[(X(t) - m_X(t))(X(s) - m_X(s))].$$

The diagonal of the covariance function, i. e., $\gamma_X(t, t)$ is called the *variance* and $\mathbb{E}[X(t)^2]$ the *second moment* which is equal to the variance if $m_X(t) = 0$.

An interesting class of stochastic processes is the one, where the mean function and the covariance function do not depend on where we are in time. These processes have a constant mean in time and a covariance that just depends on the distance of the two points in time. More formally, let us define the following:

Definition 1.3. A stochastic process X is called *stationary* if it satisfies for all $n \in \mathbb{N}$, $h \in \mathbb{R}$, and a finite collection of time points $t_1, \ldots, t_n \in \mathbb{T}$ satisfying $t_1 + h, \ldots, t_n + h \in \mathbb{T}$ that

$$(X(t_1), \ldots, X(t_n)) \simeq (X(t_1 + h), \ldots, X(t_n + h)),$$

where $\simeq$ denotes equality in distribution. This is also known as *strong stationarity*.

Unfortunately, we are often not in place to check these strong conditions and are therefore happy if we have processes that satisfy a weaker form of stationarity, which we introduce in what follows.

Definition 1.4. A stochastic process X is called *weakly stationary* if it satisfies for all $s, t \in \mathbb{T}$ that

- $\gamma_X(t,t) < +\infty$,
- $m_X(t) = m_X(s)$, i. e., m_X is constant, and
- for all $h \in \mathbb{R}$ such that $s+h, t+h \in \mathbb{T}$

$$\gamma_X(s,t) = \gamma_X(s+h, t+h).$$

So far we have formulated all properties of stochastic processes with respect to any set $\mathbb{T}$. Let us in what follows give important examples of discrete-time stochastic processes.

Example 1.5. A basic very important stochastic process in discrete time is *discrete white noise* $Z : \Omega \times \mathbb{Z} \to \mathbb{R}$ which is a weakly stationary process that satisfies

- $m_Z(t) = 0$ for all $t \in \mathbb{Z}$ and
- $\gamma_Z(s,t) = \sigma^2 \mathbb{1}_{\{t\}}(s)$ for all $s, t \in \mathbb{Z}$, i. e., $\gamma_Z(s,t) = 0$ for $s \neq t$ and $\gamma_Z(t,t) = \sigma^2$.

Especially in time series analysis, the class of so-called ARMA processes contains typical examples of linear processes in discrete time. We introduce a specific example below.

Example 1.6. An ARMA$(1,1)$ *process* $X : \Omega \times \mathbb{Z} \to \mathbb{R}$ is a stochastic process that satisfies for some $\phi, \theta \in \mathbb{R}$ and a white noise process Z the recursion formula

$$X(t) - \phi X(t-1) = Z(t) + \theta Z(t-1).$$

The probably most well-known example of a stochastic process in continuous time is Brownian motion which we will discuss in detail in the next section.

2. Brownian motion

Let us consider $\mathbb{T} = \mathbb{R}_+$ or $\mathbb{T} = [0, T]$ for some finite $T > 0$.

Definition 2.1. A *Brownian motion* $W : \Omega \times \mathbb{T} \to \mathbb{R}$ is a stochastic process that satisfies

- $W(0) = 0$,
- W is P-almost surely continuous,
- W has independent increments,
- $W(t) - W(s) \sim \mathcal{N}(0, t-s)$ for all $0 \leq s \leq t$.

As a mathematical object, Brownian motion is also known as *Wiener process*.

For $\mathbb{T} = [0, T]$ let us consider a partition of time into N intervals given by

$$0 = t_0 < t_1 < \cdots < t_N = T.$$

Typical partitions are equidistant partitions such that $h := N^{-1}$ and

$$t_n = n \cdot h.$$

In order to obtain nested grids of time discretizations which will be of help in the simulations in this project, we will typically set $N := 2^i$ for some $i \in \mathbb{N}$.

From the definition of Brownian motion we obtain that Brownian motion can be rewritten as a telescopic sum

$$W(T) = \sum_{n=1}^{N} W(t_n) - W(t_{n-1}),$$

where all *increments* $W(t_n) - W(t_{n-1})$ are independent $\mathcal{N}(0,h)$-distributed random variables. Therefore we can simulate the path of a Brownian by simulating in each time step a random number $\eta \sim \mathcal{N}(0,h)$ and setting

$$W(t_n) := W(t_{n-1}) + \eta.$$

In order to generate the same Brownian sample path on different resolutions, we observe that for grids with step size $N_1 := 2^i$ and $N_2 := 2^{i+1}$, the first two increments on the fine grid of the Brownian motion are summed together the increment of the Brownian motion on the coarser grid. More specifically, if η_1 and η_2 are the increments on the grid with respect to N_2 time steps, then

$$\tilde{\eta}_1 := \eta_1 + \eta_2$$

is the first increment of the same sample of the Brownian motion on the coarser grid with N_1 grid points. This observation extends to all grid points on the time interval and to all neighboring resolutions of the time grid.

The first task of the project is to simulate a Brownian sample path on a sequence of different time grids.

3. Approximation of stochastic processes

Let us in this section assume that $\mathbb{T} = [0, T]$, i. e., that we are working on a finite time interval, which is reasonable for computations in finite time. We are interested in simulating the stochastic process $X : \Omega \times \mathbb{T} \to \mathbb{R}$ given by

$$X(t) = \exp\left(\left(\mu - \frac{\sigma^2}{2}\right) t + \sigma W(t)\right),$$

where W denotes a Brownian motion. This example is motivated from the theory and approximation of stochastic differential equations. Although we know the process explicitly in this specific case, we will try to approximate it by a discrete time process and show how it converges to the theoretical process. We do this for simplicity in order to really control the errors explicitly. In applications, we will usually not know the process X explicitly but only have an implicit description of it. In the given case, we can argue that it makes sense to approximate the process since the computation of the exponential function is expensive if we need it at many times t compared to other operations on a computer.

Instead of computing X, let us compute $X_h : \Omega \times \{t_n := n \cdot h, n = 0, \ldots, N\} \to \mathbb{R}$ by the recursion scheme

$$X_h(t_n) = (1 + h\mu)X_h(t_{n-1}) + \sigma X_h(t_{n-1})(W(t_n) - W(t_{n-1}))$$

and $X_h(0) = 1$.

If we simulate sample paths of the true process and its approximation, we see in pictures that the difference gets smaller for finer time grids. To quantify the error mathematically, let us introduce two different concepts of errors.

Definition 3.1. Let X be a stochastic process and $(X_h, h \in (0, 1])$ be a family of approximations of the process. Then

$$\mathbb{E}\left[(X(T) - X_h(T))^2\right]^{1/2}$$

is called the *strong error*, also known as L^2 or root mean squared error.

The family $(X_h, h \in (0, 1])$ is said to *converge strongly* to X if

$$\lim_{h \to 0} \mathbb{E}\left[(X(T) - X_h(T))^2\right]^{1/2} = 0.$$

It converges with *rate* γ if there exists a constant C such that for all h sufficiently small

$$\mathbb{E}\left[(X(T) - X_h(T))^2\right]^{1/2} \leq Ch^\gamma.$$

Since we are usually not able to compute the expectation in this expression explicitly, one possibility is to approximate it by a Monte Carlo simulation. We therefore can approximate the strong error for sufficiently large $M \in \mathbb{N}$ by

$$\mathbb{E}\left[(X(T) - X_h(T))^2\right]^{1/2} \approx \left(\frac{1}{M}\sum_{m=1}^{M}((X(T)^{(m)} - X_h(T)^{(m)})^2\right)^{1/2},$$

which will be one of the tasks of this project.

Let us continue with a second type of convergence which is of relevance if we are interested in the convergence of so-called *quantities of interest* related to the distribution of the process instead of properties of the sample paths.

Definition 3.2. Let X be a stochastic process and $(X_h, h \in (0,1])$ be a family of approximations of the process. Furthermore let $\phi : \mathbb{R} \to \mathbb{R}$ be a function within a suitable class of test functions. Then

$$|\mathbb{E}[\phi(X(T))] - \mathbb{E}[\phi(X_h(T))]|$$

is called the *weak error*.

The family $(X_h, h \in (0,1])$ is said to *converge weakly* to X if

$$\lim_{h\to 0} |\mathbb{E}[\phi(X(T))] - \mathbb{E}[\phi(X_h(T))]| = 0$$

for all test functions ϕ. It converges with *rate* γ if there exists a constant C such that for all h sufficiently small

$$|\mathbb{E}[\phi(X(T))] - \mathbb{E}[\phi(X_h(T))]| \leq Ch^\gamma.$$

For weak errors, if we assume that we know $\mathbb{E}[\phi(X(T))]$ explicitly, the computation of the error can be done using the Monte Carlo estimator

$$|\mathbb{E}[\phi(X(T))] - \mathbb{E}[\phi(X_h(T))]| \approx \left|\mathbb{E}[\phi(X(T))] - \frac{1}{M}\sum_{m=1}^{M}\phi(X_h(T))^{(m)}\right|$$

for a sufficiently large number of sample paths M.

As has been seen in the lecture and the previous project, the estimation of the error is very sensitive to the choice of M. Be aware of that in your simulations. You should especially keep in mind that the total error behaves additive, i. e., is given by

$$\frac{1}{\sqrt{M}} + h^\gamma.$$

This means that whenever h is very small, the Monte Carlo error will dominate and be visible in your simulated convergence rates.

We should mention that the weak error is bounded by the strong error if ϕ is Lipschitz continuous. This leads to the same rate of weak convergence as obtained for strong convergence. In this project we will observe that this is not optimal and a general rule of thumb holds that says that the weak rate of convergence is twice the strong one.

4. Questions

Consider the stochastic process $X : \Omega \times [0,1] \to \mathbb{R}$ given by

$$X(t) = \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W(t)\right)$$

and its approximation given by the recursion formula

$$X_h(t_n) = (1 + h\mu)X_h(t_{n-1}) + \sigma X_h(t_{n-1})(W(t_n) - W(t_{n-1}))$$

with $X_h(0) = 1$. Consider further the sequence of time discretizations with

$$h_i = 2^{-i}$$

for $i = 1, \ldots, 10$ (or larger if your computer allows for that). Let us set $\mu := \sigma := 1$, but you are welcome to vary the parameters as you like.

Program all tasks in Python.

1. (1 pt). Compute a sample path of a Brownian motion at all resolutions $(h_i, i = 1, \ldots, 10)$ based on the same noise. Plot your result and explain it.

2. (1 pts). Compute a sample path of X and sample paths $(X_{h_i}, i = 1, \ldots, 10)$ based on the same noise. Plot all results in the same graph. Describe your observations.

3. (1.5 pts). Estimate the strong error with a Monte Carlo simulation based on $M = 1000$ (or more if you like) for all $(h_i, i = 1, \ldots, 10)$. Plot your results with h in the x-axis and the strong error in the y-axis in a loglog plot and add a reference slope $h^{1/2}$. Describe your findings.

4. (1.5 pts). Estimate the weak error of

$$|\mathbb{E}[X(1)] - \mathbb{E}[X_h(1)]|$$

with a Monte Carlo simulation based on $M = 1000$ (and as many more as you can compute in a reasonable time) for all $(h_i, i = 1, \ldots, 10)$. Observe that

$$\mathbb{E}[X(1)] = \exp(\mu).$$

Plot your results with h in the x-axis and the weak error in the y-axis in a loglog plot and add a reference slope h. Describe your findings.

5. (2 pts). Do the same as in *4.* but choose another test function than $\phi = \mathrm{Id}$ or any linear affine transformation of it. Plot and describe your results.

Assignment A6 has a maximum of *8 points* and must be written as a single LaTeXreport and be submitted to Canvas. Notice the "recommended deadline" on the course webpage.

Please use the recommended report template provided on the course page, and in case you worked out the exercises with some student, remember to write the name of said student as a footnote in the report front page. Also, recall that each submitted report and the code therein is an *INDIVIDUAL* submission, not group work.

Finally, since you have to write a proper report: as from the provided template, you are also asked to produce some background on the methodology you use. So do not just write answers to the exercise questions.

For a given project report, 0.5 points will be deducted if the report is not clearly structured or is otherwise hard to understand. Likewise, 0.5 points will be deducted if the code attached to the report is not properly structured and commented. The report should not be longer than 10 pages including figures, but excluding appendix. Figures and axes labels should be big enough to be readable if printed. It is OK to use colors.

Full details on grading are of course available at `https://chalmers.instructure.com/courses/10351/pages/coursepm`