Advanced computational methods-Lecture 2

1 Brief Introduction to martingales

Below, it will be assumed that X is a real valued process, adapted to some given filtration $\{\mathcal{F}_t\}$ with $\mathbb{E}|X_t| < \infty$.

Definition 1. *The process $X = \{X_t\}$ adapted to $\{\mathcal{F}_t\}$ is called a submartingale (resp. supermartingale) if for all $0 \le s < t < \infty$, $\mathbb{E}(X_t|\mathcal{F}_s) \ge X_s$ a.s. (resp. $\mathbb{E}(X_t|\mathcal{F}_s) \le X_s$) The process is called a* **martingale** *if it is both a submartingale and a supermartingale, i.e.*

$$\mathbb{E}(X_t|\mathcal{F}_s) = X_s, \ a.s.$$

Proposition 1. *Suppose the filtration $\{\mathcal{F}_t\}$ is right continuous, and $\mathcal{F}_0$ contains all measure zero sets in $\mathcal{F}$. Let X be a submartingale such that $t \mapsto \mathbb{E}X_t$ is right-continuous. Then, there is a modification of X which is cadlag (right continuous with left limits).*

Jensen's inequality tells us that if X is a submartingale, then for any convex function, $\varphi(X)$ is a submartingale if $\varphi(X)$ is integrable.

1.1 Optional stopping

Optional stopping basically wants to investigate what would happen if we stop a martingale at stopping times. The motivation is from gambling games. A natural question is whether one can choose the time to quit in a fair game to gain fortune. The answer basically is 'no', and this is what optional stopping studies.

Theorem 1. *Let X be a submartingale with right continuous paths and σ, τ be two stopping times. Then, for any $T > 0$, one has*

$$\mathbb{E}[X_{\tau\wedge T}|\mathcal{F}_\sigma] \ge M_{\sigma\wedge\tau\wedge T}.$$

From this theorem, it is clear that for a martingale, we have the equality holds. Hence, one has

Corollary 1. *Let M be right continuous martingale. The stopped process $M^\tau := \{M_{\tau\wedge t}\}$ is also a martingale. If M is square integrable, so is M^τ.*

This corollary basically says there is no strategy to gain fortune in a fair game.

In general, if we drop the T variable, the conclusion no longer holds. Thinking about the example in the previous lecture, $M_n := -X_n$ is a martingale and .. The reason is that M_n does not converge in L^1.

However, if the martingale has a last element, i.e. there exists M_∞ such that $M_t \to M_\infty$ a.s., and $\mathbb{E}(M_\infty | \mathcal{F}_t) = M_t$ (which is equivalent to $M_t \to M_\infty$ in $L^1(\mathbb{P})$), then one can expect the equality to hold.

Proposition 2. *If $X = \{X_t\}$ is a submartingale with last element X_∞. Then, for any two stopping times $\sigma \leq \tau$, one has*

$$\mathbb{E}(X_\tau | \mathcal{F}_\sigma) \geq X_\sigma.$$

The equality holds when X is a martingale with a last element.

In the example previously, $X_\tau = 1$. If we take $\sigma = t$, clearly, the equality does not hold.

1.2 Important inequalities and martingale convergence theorem

The first inequality is to control the maximum of the martingale.

Theorem 2. *Let X be a right continuous submartingale. Then, for any $r > 0$ and $T > 0$:*

$$\mathbb{P}\left(\sup_{0 \leq t \leq T} X_t \geq r\right) \leq r^{-1} \mathbb{E}[X_T^+],$$

and

$$\mathbb{P}\left(\inf_{0 \leq t \leq T} X_t \leq -r\right) \leq r^{-1} (\mathbb{E}[X_T^+] - \mathbb{E}[X_0]).$$

The Doob's inequality is more useful:

Theorem 3. *Let X be a nonnegative right continuous submartingale. Then, for any $p \in (1, \infty)$:*

$$\mathbb{E}\left[(\sup_{0 \leq t \leq T} X_t)^p\right] \leq (\frac{p}{p-1})^p \mathbb{E}[M_T^p].$$

Imagine that you have some martingale $M = (M_t)$. Then, $|M_t|$ and M_t^+ are submartingales. Then, for example, the moments of $\sup_{0 \leq t \leq T} |M_t|$ can be controlled by the moments of $|M_T|$ (or the moments of M_T if p is even). This is particularly useful when esitmating the stochastic integral in SDEs. Note that T can be replaced by **bounded** stopping times.

Another important inequality for martingale is the **Burkholdr-Davis-Gundy** inequalities (see Theorem 3.28 in the book by Karatzas and Shreve). It uses the quadratic variation of a martingale to control the moments of $\sup_{0\le t\le T}|M_t|$. We omit it here. If you are interested, you may read.

By estimation of the so-called 'upcrossing', one can have the following martingale convergence theorem (see Theorem 3.15 in the book by Karatzas and Shreve):

Theorem 4. *Let X be a right-continuous submartingale with $\sup_{t\ge 0}\mathbb{E}X_t^+ < \infty$. Then, X_t converges almost surely to some X_∞ a.s. with $\mathbb{E}|X_\infty| < \infty$.*

Intuitively, this is like the basic claim in calculus: an increasing sequence with an upper bound has a limit. This martingale convergence theorem is also not quite related to stochastic integration and our course.

1.3 Doob-Meyer decomposition

The intuition of this decomposition is that if you subtract kind of "mean" from a submartingale, the remaining part will be fluctuation with mean zero, and then it is like a martingale. This intuition then leads to the Doob-Meyer decomposition.

Definition 2. *We say X is of class DL if for any $u \in (0,\infty)$, the random variables $\{X_\tau : \tau \text{ is a stopping time with } \tau \le u\}$ is uniformly integrable.*

Lemma 1. *A right-continuous nonnegative submartingale is of class DL. In particular, if M is a square integrable martingale, then M^2 if of calss DL.*

The Doob-Meyer decomposition says

Theorem 5. *Let the filtration be right continuous and complete. Let X be a right continuous submartingale of class DL. Then, there is a* **unique** *nondecreasing predictable process A such that $X - A$ is a martingale.*

Predictable means $\{X_t \le t\} \in \mathcal{F}_{t-}$. Intuitively, this says one can predict the near future using current information.(This does not say the process can then be determined totally since the information in the filtration is not solely given by the process.) In many texts, people say "A is increasing and natural". The definition of *natural process* is given in these books. One can show that if the process is nondecreasing, the being predictable is the same as being natural.

1.4 A metric space for martingales

Often we consider the space of square integrable martingales, $\mathcal{M}_2$:

$$\|M\|_{\mathcal{M}_2} = \sum_{k=1}^{\infty} 2^{-k}(1 \wedge \|M_k\|_{L^2(\mathbb{P})}).$$

It can be verified easily that

$$(\mathcal{M}_2, \|\cdot\|_{\mathcal{M}_2})$$

is a metric space. Due to $1 \wedge \cdot$, this is not a Banach space.

Note that if you remove $1 \wedge \cdot$, then the space cannot contain all square integrable martingales, but then the quantity $\|\cdot\|$ becomes a norm.

2 Quadratic variation

2.1 quadratic variation with right continuous paths

Definition 3. *Given a process Y, the quadratic variation $[Y]$ is a stochastic process such that $t \mapsto [Y](t,\omega)$ is nondecreasing for all ω and*

$$[Y](t,\omega) = \lim_{mesh \to 0} \sum_{i}^{n} (Y_{t_{i+1}} - Y_{t_i})^2, \textit{ in probability}$$

Exercise: Prove that if $X(t)$ is a continuously differentiable, then $[X] = 0$. Prove that the quadratic variation is given by $[W] = t$.

Since Brownian motion has rough paths, the quadratic variation is nonzero. In fact,

$$[W] = t.$$

This formally means

$$(dW)^2 = dt.$$

This explains why in usual ODEs, we only have $df(X) = f'(X)dX$ but for diffusion processes, $df(W)$ contains other terms besides $f'(W)dW$.

Theorem 6. *Let M be a right continuous martingale. The quadratic variation $[M] = ([M]_t)$ exists and there is a version such that the paths are non-decreasing, and right continuous, adapted to the underlying filtration with $[M]_0 = 0$. If M is continuous, $[M]$ is also continuous.*

Exercise: show that $\mathbb{E}[M]_t = M_t^2 - M_0^2$.

In fact,

Proposition 3. *If M is a right continuous L^2 martingale, then $M_t^2 - [M]_t$ is a martingale.*

The quadratic covariation is

$$[X, Y] = [\frac{1}{2}(X + Y)] - [\frac{1}{2}(X - Y)]$$

Intuitively, this is to define

$$\sum_i (X_{t_{i+1}} - X_{t_i})(Y_{t_{i+1}} - Y_{t_i})$$

The previous definition is used due to some path properities.

Hence, formally, we have

$$d[X, Y] = dXdY.$$

Remark 1. *As we have seen, the quadratic variation is introduced because $(dW)^2$ has nontrivial contribution. In general, if we consider diffusion processes driven by Brownian motions, we need to consider quadratic variation and the following product rule holds*

$$d(XY) = YdX + XdY + d[X, Y]$$

Here, YdX, XdY should be understood in Itô sense as we shall see soon.

2.2 Predictable quadratic variation

Using the Doob-Meyer decomposition, one may find another option to define quadratic variation. In fact, M^2 is right continuous submartingale, so it is of *class DL*. Hence, by Doob-Meyer decomposition we can write

$$M^2 = \langle M \rangle + (M^2 - \langle M \rangle),$$

such that $M^2 - \langle M \rangle$ is a martingale. The process $\langle M \rangle$ is predictable, non-decreasing, and unique, called the predictable quadratic variation.

In general, $[M]$ is not the same as $\langle M \rangle$ because the former is right-continuous while the latter is predictable. However, if the process M is continuous, $[M]$ is also continuous, and thus predictable. Then, $[M] = \langle M \rangle$.

3 Stochastic integration

3.1 Stochastic integration with respect to Brownian motions

Suppose we want to define $\int_0^t WdW$. Consider the Riemann sum

$$S(\pi) = \sum_i W_{s_i}(W_{t_{i+1}} - W_{t_i})$$

where

$$s_i = (1-u)t_i + ut_{i+1}.$$

The computation starts with the following algebra identity

$$b(a-c) = \frac{1}{2}(a^2 - c^2) - \frac{1}{2}(a-c)^2 + (b-c)^2 + (a-b)(b-c)$$

Taking the sum, the first term is simply $\frac{1}{2}W_t^2$. The second term is the quadratic variation. The third term, by similar computation of quadratic variation, we can show that it converges in L^2 to ut:

$$\mathbb{E}(\sum_i (W_{s_i} - W_{t_i})^2) = \sum_i u(t_{i+1} - t_i) = ut,$$

$$Var(\sum_i (W_{s_i} - W_{t_i})^2) = \sum_i Var((W_{s_i} - W_{t_i})^2) = \sum_i 2(s_i - t_i)^2 \leq 2tmesh(\pi) \to 0.$$

The third term has mean zero and variance converging to zero. Hence, the L^2 limit is given by

$$\frac{1}{2}W_t^2 - \frac{1}{2}t + ut.$$

We conclude the following

- The limit of the Riemann sum depends on the choice of sample point!
- If $u = 1/2$, we have the chain rule. However, $\frac{1}{2}W_t^2$ is not a martingale.
- If $u = 0$, we do not have chain rule but $\frac{1}{2}W_t^2 - \frac{t}{2}$ is a martingale.

If we choose the midpoint as the sample point, we get the Stratonovich integral. If we use the left point, the resulted integral is the Itô integral. Since the Itô integrals give martingales

Remark 2. *Show that if F has bounded total variation, then*

$$\lim_{mesh(\pi)\to 0} \sum_i F_{s_i}(F_{t_{i+1}} - F_{t_i})$$

is independent of the choice of u.

Remark 3. *For the stochastic integral with respect to cadlag semimartingales, the integrand should be predictable to make sense.*

Rigorous definition of Itô integral

The rigorous stochastic integral can be established for $X \in L^2([0,T]\times\Omega)$, i.e.

$$\|X\|^2_{L^2([0,T]\times\Omega)} := \mathbb{E}\int_{[0,T]} |X(t,\omega)|^2 dt < \infty.$$

- First of all, for the simple predictable process

  $$X_t(\omega) = \eta_0(\omega)1_{(0)}(t) + \sum_{i=1}^{n-1} \xi_i(\omega)1_{(t_i,t_{i+1}]}(t)$$

  where $0 = t_1 < t_2 < \ldots < t_n$. Predictable means that the state at t can be referred by the information in $s < t$. Also, ξ_i should be square integrable. The stochastic integral is defined by

  $$(X \cdot B)(t) = \int_0^t X dB = \sum_{i=1}^{n-1} \xi_i(\omega)(B_{t_{i+1}\wedge t}(\omega) - B_{t_i \wedge t}(\omega))$$

  Clearly, $X \cdot B$ is a martingale in $\mathcal{M}_2$ and we have the Itô isometry:

  $$\mathbb{E}[(X \cdot B)_t^2] = \mathbb{E}\int_0^t X_s^2\, ds, \quad \forall t \geq 0.$$

- The significant fact of the Itô isometry is that if $\{X_n\}$ is a sequence of simple predictable processes and $X_n \to X$ in $L^2([0,T]\times\Omega)$, then $X_n \cdot B$ is a Cauchy sequence in $\mathcal{M}_2$. Then, the limit is defined to be $(X \cdot B)(t) = \int_0^t X dB$.

  This is good enough because any $X \in L^2([0,T]\times\Omega)$ can be approximated by simple predictable processes.

  In the example above, $\sum_i B_{t_i} 1_{(t_i,t_{i+1}]}$ approximates B, and this corresponds to $u = 0$.

To summarize, for $X \in L^2([0,T]\times\Omega)$, we can define $X \cdot B = \int_0^t X dB$ which is a martingale and the Itô isometry holds

$$\mathbb{E}\left(\int_0^t X dB\right)^2 = \mathbb{E}\int_0^t X_s^2\, ds$$

There are extensions to processes that are not in L^2. Those who are interested can read the reference.

3.2 Stochastic integration with respect to continuous square integrable martingales

Since we will consider SDEs like $dX = b(X)dt + \sigma(X)dW$ and consider increment of the form $df(X)$. This will leads to integration with respect to X. Hence, it is necessary to consider integration with respect to general continuous martingale. For martingales with possible jumps, see the remark below or the discussion in Timo's notes.

Now, for a general continuous square integrable martingale, the space of integrands becomes $L^2([0,T]\times\Omega, \mu_M)$ with the measure μ_M on $([0,T]\times\Omega, \mathcal{R})$ defined by

$$\mu_M(A) = \mathbb{E}\int_{[0,\infty)} 1_A(t,\omega)d[M]_t.$$

Here, $\mathcal{R}$ is the so-called predictable sigma-field. In fact, for M to be continous, one can enlarge this to general $\mathscr{B}([0,T]\times\Omega)$. However, if M has jumps, the integral defined this way for non-predictable will include the effects of jumps (see Remark 4 below).

As soon a progressively measurable process is in this space, one can use the simple processes $\sum_i \xi_i 1_{(t_i,t_{i+1}]}$ with $\xi_i \in \mathcal{F}_{t_i}$ to approximate. Then, the stochastic integrals can be defined similarly.

The important properties include

- $X\cdot M := \int_0^t XdM$ is a square integrable martingale, continuous.
- The isometry holds

$$\mathbb{E}((X\cdot M)_t)^2 = \int_{[0,t]\times\Omega} X^2 d\mu_M = \mathbb{E}\int_{[0,t]} X^2 d[M]_t.$$

Remark 4. *Note that the simple processes we consider here are all predictable. The reason to use $1_{(t_i,t_{i+1}]}$ is that $\int 1_{(t_i,t_{i+1}]}dM = M_{t_{i+1}} - M_{t_i}$. If we use $1_{[t_i,t_{i+1})}$, then we have $M_{t_{i+1}^-} - M_{t_i^-}$.*

If M has jumps, one should be careful with the integrands. In fact, for M with jumps, the simple processes of the form $\sum_i \xi_i 1_{(t_i,t_{i+1}]}$ will converge to the predictable processes only under the metric corresponding to M. If one uses $\sum_i \xi_i 1_{[t_i,t_{i+1})}$, it can approximate general processes with jumps, but the integral may not be a Martingale. For example, let $X = M = N = \{N_t\}$ be the Poisson process. Then, then $\int\sum_i N(t_i)1_{(t_i,t_{i+1}]}dN \to \int N(t^-)dN$ which is a martingale, but $\int\sum_i N(t_i)1_{[t_i,t_{i+1})}dN \to \int NdN = \int N(t^-)dN + \sum(\Delta N)^2$. It is a martingale plus a non-decreasing process, which is no longer a martingale. Due to this reason, when M has jumps, the integrand should be predictable.