Project 1

I suggest that you bring a copy of this sheet to the Friday session and make rough notes on it while I explain some of the ideas. You should then go home and write a reasonably self-contained account in your notebook. You may consult any texts you wish and you may ask me or anyone else as many questions as you like.

Please do not just copy out standard proofs without understanding. Please do not just copy from someone else's notebook.

In your weekly session—DON'T FORGET TO ARRANGE A TIME WITH ME—*I will discuss with you any difficulties you have with producing an account in your own words. I will also point out refinements, if you are interested.*

At the end of the semester, I will look at your notebook to make up a grade. By that time, you should have a pretty good written account of a significant chunk of stochastic calculus.

Things to explain in your notebook:

(i) filtrations and stochastic processes adapted to a filtration
(ii) stopping times and related sigma-fields
(iii) (sub/super)martingales in continuous time
(iv) How does progressive measurability help?
(v) cadlag sample paths
(vi) versions of stochastic processes
(vii) standard filtrations: Why are they convenient?
(viii) cadlag versions of martingales adapted to standard filtrations

Please pardon my grammar. This sheet is witten in note form, not in real sentences.

Filtrations and stochastic processes

Fixed probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Negligible sets $\mathcal{N} := \{N \in \mathcal{F} : \mathbb{P}N = 0\}$. Without loss of generality the probability space is complete.

completeness needed later

Time set $T \subseteq \mathbb{R} \cup -\infty \cup \{\infty\}$. Filtration $\{\mathcal{F}_t : t \in T\}$: set of sub-sigma-fields of $\mathcal{F}$ with $\mathcal{F}_s \subseteq \mathcal{F}_t$ if $s < t$. Think of $\mathcal{F}_t$ as "information available at time t"?

If $\infty \notin T$ define $\mathcal{F}_\infty := \sigma\left(\cup_{t\in T} \mathcal{F}_t\right)$.

Stochastic process $\{X_t : t \in T\}$: a set of $\mathcal{F}$-measurable random variables. Write $X_t(\omega)$ or $X(t, \omega)$. We can think of X as a map from $T \times \Omega$ into $\mathbb{R}$.

Say that X is ***adapted to the filtration*** if X_t is $\mathcal{F}_t$-measurable for each $t \in T$. Value of $X_t(\omega)$ can be determined by the "information available at time t".

Stopping times

Function $\tau : \Omega \to \overline{T} := T \cup \{\infty\}$ such that $\{\omega : \tau(\omega) \le t\} \in \mathcal{F}_t$ for each $t \in T$. Check that τ is $\mathcal{F}_\infty$-measurable. Define

$$\mathcal{F}_\tau := \{F \in \mathcal{F}_\infty : F\{\tau \le t\} \in \mathcal{F}_t \text{ for each } t \in T\}$$

Check that $\mathcal{F}_\tau$ is a sigma-field. Show that τ is $\mathcal{F}_\tau$-measurable. Show that an $\mathcal{F}_\infty$-measurable random variable Z is $\mathcal{F}_\tau$-measurable if and only if $Z\{\tau \le t\}$ is $\mathcal{F}_t$-measurable for each $t \in T$.

Progressive measurability

Problem: If $\{X_t : t \in T\}$ is adapted and τ is a stopping time, when is the function

$$\omega \mapsto X(\tau(\omega), \omega)\{\tau(\omega) < \infty\}$$

$\mathcal{F}\tau$-measurable? Perhaps simplest to think only of the case where $T = \mathbb{R}^+$, equipped with its Borel sigma-field $\mathcal{B}(T)$.

Warmup: Suppose τ takes values in T. Show $X(\tau(\omega), \omega)$ is $\mathcal{F}$-measurable.

$$\begin{array}{ccccc} \omega & \mapsto & (\tau(\omega), \omega) & \mapsto & X(\tau(\omega), \omega) \\ \Omega & & T \times \Omega & & \mathbb{R} \\ \mathcal{F} & & \mathcal{B}(T) \otimes \mathcal{F} & & \mathcal{B}(\mathbb{R}) \end{array}$$

If X is $\mathcal{B}(T) \otimes \mathcal{F}\backslash\mathcal{B}(\mathbb{R})$-measurable, we get $\mathcal{F}\backslash\mathcal{B}(\mathbb{R})$-measurability for the composition.

Abbreviate $\mathcal{B}([0,t])$, the Borel sigma-field on $[0,t]$, to $\mathcal{B}_t$.

- Now suppose X is ***progressively measurable***, that is, the restriction of X to $[0,t] \times \Omega$ is $\mathcal{B}_t \otimes \mathcal{F}_t$-measurable for each $t \in T$. For a fixed t, write Y for the restriction of X to $[0,t] \times \Omega$. Show that

$$X(\tau(\omega), \omega)\{\tau(\omega) \le t\} = Y(\tau(\omega) \wedge t, \omega)\{\tau(\omega) \le t\}$$

  Adapt the warmup argument to prove that $Y(\tau(\omega) \wedge t, \omega)$ is $\mathcal{F}_\tau$-measurable. Then what? Conclude that $X(\tau(\omega), \omega)\{\tau(\omega) < \infty\}$ is $\mathcal{F}_\tau$-measurable.

- Show that an adapted process with right-continuous sample paths is progressively measurable. Argue as follows, for a fixed t. Define $t_i = it/n$ and

$$X_n(s, \omega) = X(0, \omega) + \sum_{i \le n} X(t_i, \omega)\{t_{i-1} < s \le t_i\} \qquad \text{for } 0 \le s \le t.$$

  Show that X_n is $\mathcal{B}_t \otimes \mathcal{F}_t$-measurable and X_n converges pointwise to the restriction of X to $[0,t] \times \Omega$.

Cadlag

Define $\mathbb{D}(T)$ as the set of real valued functions on T that are right continuous and have left limits at each t. (Modify the requirements suitably at points not in the interior of T.) Say that functions in $\mathbb{D}(T)$ are cadlag on T.

Say that a process X has ***cadlag sample paths*** if the function $t \mapsto X(t, \omega)$ is cadlag for each fixed ω.

A typical sample path problem

For a fixed integable random variable ξ, define $X_t(\omega) = \mathbb{P}(\xi \mid \mathcal{F}_t)$. Note that $\{(X_t, \mathcal{F}_t) : t \in T\}$ is a martingale. Remember that each X_t is defined only up to an almost sure equivalence. Question: Must X be progressively measurable?

To keep things simple, assume $T = [0,1]$.

Suppose Y_t is another choice for $\mathbb{P}(\xi \mid \mathcal{F}_t)$. Note that

$$N_t := \{\omega : X_t(\omega) \ne Y_t(\omega)\} \in \mathcal{N}$$

That is, the stochastic process Y is a ***version*** of X. However, the sample paths of X and Y can be very different:

$$\{\omega : X(\cdot, \omega) \ne Y(\cdot, \omega)\} \subseteq \cup_{t \in T} N_t$$

A union of uncountably many negligible sets need not be negligible.

We need to be careful about the choice of the random variable from the equivalence class corresponding to $\mathbb{P}(\xi \mid \mathcal{F}_t)$.

How to construct a cadlag version of X

Without loss of generality (why?) suppose $\xi \geq 0$.

- First build a nice "dense skeleton". Define $S_k := \{i/2^k : i = 0, 1, \dots, 2^k\}$ and $S = \cup_{k\in\mathbb{N}} S_k$. For each s in S, choose arbitrarily a random variable X_s from the equivalence class $\mathbb{P}(\xi \mid \mathcal{F}_s)$.
- Show that
$$\mathbb{P}\{\max_{s\in S_k} X_s > x\} \leq \mathbb{P}X_0/x \qquad \text{for each } x > 0.$$
Let k tend to infinity then x tend to infinity to deduce that $\sup_{s\in S} X_s < \infty$ almost surely.
- For fixed rational numbers $0 < \alpha < \beta$, invoke Dubin's inequality to show that the event
$$A(\alpha, \beta, k, n)$$
$$:= \{\text{the process } \{X_s : s \in S_k\} \text{ makes at least } n \text{ upcrossings of } [a, \beta]\ \}$$
has probability less than $(\alpha/\beta)^n$.
- Let k tend to infinity, then n tend to infinity, then take a union over rational pairs to deduce existence of an $N \in \mathcal{N}$ such that, for $\omega \in N^c$, the sample path $X(\cdot, \omega)$ (as a function on S) is bounded and
$$X(\cdot, \omega) \text{ makes only finitely many upcrossings of each rational interval .}$$
- Deduce that $\widetilde{X}_t(\omega) := \lim_{s\downarrow\downarrow t} X(s, \omega)$ exists and is finite for each $t \in [0.1)$ and each $\omega \in N^c$. Deduce also that $\lim_{s\uparrow\uparrow t} X(s, \omega)$ exists and is finite for each $t \in (0.1]$ and each $\omega \in N^c$.

$\uparrow\uparrow$ means strictly increasing and $\downarrow\downarrow$ means strictly decreasing

- Define $\widetilde{X}(\cdot, \omega) \equiv 0$ for $\omega \in N$. Show that $\widetilde{X}$ has cadlag sample paths.
- Note: $\widetilde{X}$ need not be $\mathcal{F}_t$-measurable but it is measurable with respect to the sigma-field $\widetilde{\mathcal{F}}_t := \cap_{s>t}\sigma(\mathcal{N} \cup \mathcal{F}_t)$.
- Show that $\{\widetilde{\mathcal{F}}_t : t \in [0, 1]\}$ is ***right continuous***, that is, $\widetilde{\mathcal{F}}_t = \cap_{s>t}\widetilde{\mathcal{F}}_s$, and that $\mathcal{N} \subseteq \widetilde{\mathcal{F}}_t$. [Assuming that $\mathbb{P}$ is complete, a filtration with these properties is said to be ***standard*** or to satisfy the ***usual conditions***.]
- Show that $\{(\widetilde{X}_t, \widetilde{\mathcal{F}}_t) : 0 \leq t \leq 1\}$ is a martingale with cadlag sample paths.
- Is it true that $\widetilde{X}$ is a version of X?

To complete your understanding, find a filtration (which is necessarily not standard) for which there is a martingale that does not have a version with cadlag sample paths.

Why do you think that most authors prefer to assume the usual conditions?

Small exercise on measurability

Suppose X is a bounded random variable and that $\mathcal{G}$ is a sub-sigma-field of $\mathcal{F}$. Suppose that for each $\epsilon > 0$ there is a finite set of $\mathcal{G}$-questions (that is, you learn the value of $\{\omega \in G_i\}$ for some sets of your choosing $G_1, \dots, G_N$ from $\mathcal{G}$) from which $X(\omega)$ can be determined up to a $\pm\epsilon$ error. Show that X is $\mathcal{G}$-measurable. [This problem might help you think about measurability. Imagine that you are allowed to ask the value of $\{\omega \in F\}$, but I will answer only if F is a set from $\mathcal{G}$.]

Project 2

Things to explain in your notebook:

(i) how to prove that a first passage time is a stopping time

(ii) how the Stopping Time Lemma extends from discrete to continuous time

(iii) Give some illuminating examples.

Fixed probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Negligible sets $\mathcal{N} := \{N \in \mathcal{F} : \mathbb{P}N = 0\}$. Assume the probability space is complete, that is, for all $A \subseteq \Omega$, if $A \subseteq N \in \mathcal{N}$ then $A \in \mathcal{N}$. From now on, unless indicated otherwise, also assume that all filtrations $\{\mathcal{F}_t : t \in T\}$ are standard, that is, $\mathcal{N} \subseteq \mathcal{F}_t = \mathcal{F}_{t+}$ for each t.

First passage times (a.k.a. debuts)

Suppose $\{X_t : t \in \mathbb{R}^+\}$ is adapted and that $B \in \mathcal{B}(\mathbb{R})$. Define

$$\tau(\omega) = \inf\{t \in \mathbb{R}^+ : X(t, \omega) \in B\}.$$

As usual, $\inf \emptyset := +\infty$.

- *Easy case: B open and X has right-continuous paths*
  Let S be a countable, dense subset of $\mathbb{R}^+$. Show that
  $$\{\omega : \tau(\omega) < t\} = \cup_{t>s\in S}\{X_s(\omega) \in B\} \in \mathcal{F}_t$$
  Deduce that $\{\tau \le t\} \in \mathcal{F}_{t+} = \mathcal{F}_t$.
- *Slightly harder case: B closed and X has continuous paths*
  Let $G_i := \{x : d(x, B) < i^{-1}\}$, an open set. Define $\tau_i = \inf\{t : X_t \in G_i\}$. Show that $\tau = \sup_i \tau_i$ so that $\{\tau \le t\} = \cap_{i\in\mathbb{N}}\{\tau_i \le t\} \in \mathcal{F}_t$.
- *General case: B any Borel set and X progressively measurable*
  See the handout on analytic sets. The idea is that the set
  $$D_t := \{(s, \omega) : s < t \text{ and } X(s, \omega) \in B\}$$
  is $\mathcal{B}_t \otimes \mathcal{F}_t$-measurable. The set $\{\tau < t\}$ is the projection of D_t onto Ω. A deep result about analytic sets asserts that the projection of D_t belongs to the $\mathbb{P}$-completion of $\mathcal{F}_t$. For a standard filtration, $\mathcal{F}_t$ is already complete. Thus $\{\tau < t\} \in \mathcal{F}_t$ and $\{\tau \le t\} \in \mathcal{F}_{t+} = \mathcal{F}_t$. That is, τ is a stopping time.

Preservation of martingale properties at stopping times

<1> **Stopping Time Lemma.** *Suppose $\{(X_t, \mathcal{F}_t) : 0 \le t \le 1\}$ is a positive supermartingale with cadlag sample paths. Suppose σ and τ are stopping times taking values in $[0, 1]$ and F is an event in $\mathcal{F}_\sigma$ for which $\sigma(\omega) \le \tau(\omega)$ when $\omega \in F$. Then $\mathbb{P}X_\sigma F \ge \mathbb{P}X_\tau F$.*

bounded stopping times

Proof: discrete case. Suppose both stopping times actually take values in a finite subset of points $t_0 < t_1 < \ldots < t_N$ in $[0, 1]$. Define $\xi_i := X(t_i) - X(t_{i-1})$. The superMG property means that

$$\mathbb{P}\xi_i F \le 0 \qquad \text{for all } F \in \mathcal{F}(t_{i-1})$$

Note that

$$X_\tau = X(t_0) + \sum\nolimits_{i=1}^N \xi_i\{t_i \le \tau\}$$
$$X_\sigma = X(t_0) + \sum\nolimits_{i=1}^N \xi_i\{t_i \le \sigma\}$$

so that

$$\mathbb{P}(X_\tau - X_\sigma)F = \sum\nolimits_{i=1}^N \mathbb{P}\left(\xi_i\{\sigma < t_i \le \tau\}F\right)$$

The last equality uses the fact that $\sigma \le \tau$ on F. Check that $\{\sigma < t_i \le \tau\}F$ is $\mathcal{F}(t_{i-1})$-measurable.

Proof: general case. For each $n \in \mathbb{N}$ define $\sigma_n = 2^{-n}\lceil 2^n\sigma\rceil$. That is,

$$\sigma_n(\omega) = 0\{\sigma(\omega) = 0\} + \sum\nolimits_{i=1}^{2^n} i/2^n\{(i-1)/2^n < \sigma(\omega) \le i/2^n\}$$

- Check that σ_n is a stopping time taking values in a finite subset of $[0,1]$. Question: If we rounded down instead of up, would we still get a stopping time? Check that $F \in \mathcal{F}(\sigma_n)$:

$$F\{\sigma_n \le i/2^n\} = F\{\sigma \le i/2^n\} \in \mathcal{F}(i/2^n).$$

  Define τ_n analogously.
- From the discrete case, deduce that

$$\mathbb{P}X(\sigma_n)F \ge \mathbb{P}X(\tau_n)F \qquad \text{for each } n$$

- Show that $\sigma_n(\omega) \downarrow \sigma(\omega)$ and $\tau_n(\omega) \downarrow \tau(\omega)$ as $n \to \infty$.

left-continuous paths wouldn't help—why not?

- Use right-continuity of the sample paths to deduce that $X(\sigma_n, \omega) \to X(\sigma, \omega)$ and $X(\tau_n, \omega) \to X(\tau, \omega)$ for each ω.
- Prove that $\{X(\sigma_n) : n \in \mathbb{N}\}$ is uniformly integrable. Write Z_n for $X(\sigma_n)$.
  (i) First show that $\mathbb{P}Z_n \uparrow c_0 \le \mathbb{P}X_0$ as $n \to \infty$.
  (ii) Choose m so that $\mathbb{P}Z_m > c_0 - \epsilon$. For $n \ge m$, show that $Z_n, Z_{n-1}, \dots, Z_m$ is a superMG.
  (iii) For constant K and $n \ge m$, show that

$$\begin{aligned}\mathbb{P}Z_n\{Z_n \ge K\} &= \mathbb{P}Z_n - \mathbb{P}Z_n\{Z_n < K\}\\ &\le c_0 - \mathbb{P}Z_m\{Z_n < K\}\\ &\le \epsilon + \mathbb{P}Z_m\{Z_n \ge K\}\end{aligned}$$

  (iv) Show that $\mathbb{P}\{Z_n \ge K\} \le c_0/K$, then complete the proof of uniform integrability.
- Prove similarly that $\{X(\tau_n) : n \in \mathbb{N}\}$ is uniformly integrable. Pass to the limit in the "discretized version" to complete the proof.

Problems = some possible examples for your notes

[1] Show that Lemma <1> also holds without the assumption that $X_t \ge 0$. Hint: Let M be a cadlag version of the martingale $\mathbb{P}\left(X_1^- \mid \mathcal{F}_t\right)$. Show that $Z_t := X_t + M_t$ is a positive superMG with cadlag paths.

[2] Suppose $\{(X_t, \mathcal{F}_t) : t \in \mathbb{R}^+\}$ is a positive supermartingale with cadlag sample paths. Suppose $\sigma_1 \le \sigma_2 \le \dots$ are stopping times taking values in $[0, \infty]$.

(i) Show that the sequence $X_{\sigma_n}\{\sigma_n < \infty\}$ is a superMG for a suitable filtration. Compare with UGMTP Problem 6.5.

(ii) If X is actually a positive martingale, is the sequence $X_{\sigma_n}\{\sigma_n < \infty\}$ also a martingale?

(iii) Same question as for part (ii), except that the σ_n all take values in $[0, 1]$.

[3] Suppose $\{(X_t, \mathcal{F}_t) : t \in [0,1]\}$ is a positive supermartingale with cadlag sample paths. Let $\mathcal{S}$ denote the set of all stopping times taking values in $[0, 1]$. Is the set of random variables $\{X_\sigma : \sigma \in \mathcal{S}\}$ uniformly integrable? It might help to track down the concept of superMGs of class [D].

Project 3

Things to explain in your notebook:

(i) How can Lévy's martingale characterization of Brownian motion be derived from a martingale central limit theorem?

(ii) How can stochastic processes with continuous sample paths be treated as random elements of a space of continuous functions

(iii) What is the strong Markov property for Brownian motion? Maybe sketch some sort of proof.

(iv) The completion of the filtration generated by a Brownian motion is standard.

Fixed probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Negligible sets $\mathcal{N} := \{N \in \mathcal{F} : \mathbb{P}N = 0\}$. Assume the probability space is complete, that is, for all $A \subseteq \Omega$, if $A \subseteq N \in \mathcal{N}$ then $A \in \mathcal{N}$. From now on, unless indicated otherwise, also assume that all filtrations $\{\mathcal{F}_t : t \in T\}$ are standard, that is, $\mathcal{N} \subseteq \mathcal{F}_t = \mathcal{F}_{t+}$ for each t.

McLeish <1> **Theorem.** *(McLeish 1974) For each n in $\mathbb{N}$ let $\{\xi_{nj} : j = 0, \dots, k_n\}$ be a martingale difference array, with respect to a filtration $\{\mathcal{F}_{nj}\}$, for which:*

(i) $\sum_j \xi_{nj}^2 \to 1$ *in probability;*

(ii) $\max_j |\xi_{nj}| \to 0$ *in probability;*

(iii) $\sup_n \mathbb{P} \max_j \xi_{nj}^2 < \infty$.

Then $\sum_j \xi_{nj} \rightsquigarrow N(0,1)$ *as* $n \to \infty$.

Remark. In the last right-hand side on line 5 page 202 a factor X_n is missing. We need the fact that $X_n = O_p(1)$ to prove that $Y_n \to 0$ in probability.

Levy <2> **Lévy's martingale characterization of Brownian motion.** *Suppose $\{X_t : 0 \le t \le 1\}$ is a martingale with continuous sample paths and $X_0 = 0$. Suppose also that $X_t^2 - t$ is a martingale. Then X is a Brownian motion.*

Rigorous proof that $X_1 \sim N(0,1)$. Use stopping times to cut the path into increments corresponding to the nth row of martingale differences in Theorem <1>. Omit the subscript n's. Take $\tau_0 = 0$ and

$$\tau_{i+1} = \min\left(n^{-1} + \tau_i, \inf\{t \ge \tau_i : |X(t) - X(\tau_i)| \ge n^{-1}\right)$$

For $j = 1, 2, \dots$ define

$$\xi_j := X(\tau_j) - X(\tau_{j-1})$$
$$\delta_j := \tau_j - \tau_{j-1}$$
$$V_j := \xi_j^2 - \delta_j$$

Write $\mathbb{P}_j(\dots)$ for $\mathbb{P}(\dots \mid \mathcal{F}(\tau_j))$.

- Check that $\mathbb{P}_{j-1}\xi_j = 0$ and $\mathbb{P}_{j-1}\left(V_j\right) = 0$, almost surely.
- Show that $\max_j |\xi_j| \le n^{-2}$ and $\max_j \delta_j \le n^{-1}$.
- Show that there exist $\{k_n\}$ such that $\mathbb{P}\{\sum_{j \le k_n} \delta_j \ne 1\} \to 0$ as $n \to \infty$.
- Show that $\mathbb{P}\sum_{j \le k_n} V_j = 0$ and

$$\mathbb{P}\left(\sum\nolimits_{j \le k_n} V_j\right)^2 = \sum\nolimits_{j \le k_n} \mathbb{P}V_j^2$$
$$\le n^{-1}\mathbb{P}\sum_{j \le k_n} |V_j|$$
$$\le n^{-1}\mathbb{P}\sum\nolimits_{j \le k_n}\left(V_j + 2\delta_j\right) \to 0 \qquad \text{as } n \to \infty.$$

- Deduce that $\sum_{j\le k_n}\xi_j^2 \to 1$ in probability.
- Deduce that $X(\tau_{k_n}) \rightsquigarrow N(0,1)$.
- Deduce that $X_1 \sim N(0,1)$.
- For enthusiasts: Extend the argument to show that X is a Brownian motion.

$\square$

Random elements of a function space

Suppose $\{X_t : t \in \mathbb{R}^+\}$ is a process with continuous sample paths. That is, for each fixed ω the sample path $X(\cdot,\omega)$ is a member of $C[0,\infty)$, the set of all continuous real functions (not necessarily bounded) on $\mathbb{R}^+$. Equip $C[0,\infty)$ with its ***cylinder sigma-field*** $\mathcal{C}$, which is defined as the smallest sigma-field on $C[0,\infty)$ for which each coordinate map π_t, for each $t \in \mathbb{R}^+$, is $\mathcal{C}\backslash\mathcal{B}(\mathbb{R})$-measurable. Then $\omega \mapsto X(\cdot,\omega)$ is an $\mathcal{F}\backslash\mathcal{C}$-measurable map from Ω into $C[0,\infty)$.

For fixed $t \in \mathbb{R}^+$, $\pi_t(x) := x(t)$ for $x \in C[0,\infty)$.

- Prove the last assertion. Note that $\pi_t X(\cdot,\omega) = X_t(\omega)$.
- Find some nontrivial examples of sets in $\mathcal{C}$. For example, is the set $\{x \in C[0,\infty) : \sup_t x(t) \le 6\}$ in $\mathcal{C}$?

The distribution of X is a probability measure defined on $\mathcal{C}$, the image of $\mathbb{P}$ under the map $\omega \mapsto X(\cdot,\omega)$. For example, for a standard Brownian motion, the distribution is called ***Wiener measure***, which I will denote by $\mathbb{W}$. In other words, if B is a standard Brownian motion, and at least if $f : C[0,\infty) \to \mathbb{R}^+$ is a $\mathcal{C}\backslash\mathcal{B}(\mathbb{R}^+)$-measurable function, then

$$\mathbb{P}^\omega f(X(\cdot,\omega)) = \mathbb{W}^x f(x).$$

Sometimes I will slip into old-fashioned terminology and call a real-valued function a ***functional*** if it is defined on a space of functions.

For each fixed $\tau \in \mathbb{R}^+$, define the stopping operator $K_\tau : C[0,\infty) \to C[0,\infty)$ by

$$(K_\tau x)(t) = x(\tau \wedge t) \qquad \text{for } t \in \mathbb{R}^+.$$

Decomposition of Brownian motion

Think of a standard Brownian motion $\{B_t : t \in \mathbb{R}^+\}$ as a random element of $C[0,\infty)$. For a fixed $\tau \in \mathbb{R}^+$, the process

$$Z(t) := B(\tau + t) - B(\tau) \qquad \text{for } t \in \mathbb{R}^+$$

is also a Brownian motion (with respect to which filtration?). Moreover the process Z, as a random element of $(C[0,\infty),\mathcal{C})$, is independent of $\mathcal{F}_\tau$. Proof? This fact can be reexpressed in several useful ways.

You might try your skills at generating-class arguments to establish some of the following. You might also give some special cases as examples.

Compare with the Brownian motion chapter of UGMTP.

Define the shift operator S_τ by

$$(S_\tau x)(t) = \begin{cases} 0 & \text{for } 0 \le t < \tau \\ x(t-\tau) & \text{for } t \ge \tau \end{cases}$$

(i) B has the same distribution as $K_\tau B + S_\tau \widetilde{B}$, where $\widetilde{B}$ is a new standard Brownian motion that is independent of B.

(ii) At least for each $\mathcal{C}$-measurable functional $h : C[0,\infty) \to \mathbb{R}^+$,

$$\mathbb{P}\left(h(B) \mid \mathcal{F}_\tau\right) = \mathbb{W}^x h(K_\tau B + S_\tau x) \qquad \text{almost surely.}$$

Notice that $K_\tau B$ is $\mathcal{F}_\tau$-measurable. It is unaffected by the integral with respect to $\mathbb{W}$.

(iii) For each $F \in \mathcal{F}_\tau$ and each h as in (ii),

$$\mathbb{P}Fh(B) = \mathbb{P}^\omega\left(\{\omega \in F\}\mathbb{W}^x h(K_\tau B(\cdot,\omega) + S_\tau x)\right)$$

(iv) At least for each $\mathcal{B}(\mathbb{R}) \otimes \mathcal{C}$-measurable map $f : \mathbb{R} \times C[0,\infty) \to \mathbb{R}^+$, and each $\mathcal{F}_\tau$-measurable random variable Y,

$$\mathbb{P}f(Y,B) = \mathbb{P}^\omega \mathbb{W}^x f(Y, K_\tau B + S_\tau x)$$

The ***strong Markov property*** for Brownian motion asserts that properties (i) to (iv) also hold for stopping times τ, provided we handle contributions from $\{\tau = \infty\}$ appropriately. For example, with f and Y as in (iv),

$$\mathbb{P}f(Y,B)\{\tau < \infty\} = \mathbb{P}^\omega \mathbb{W}^x f(Y, K_\tau B + S_\tau x)\{\tau < \infty\}$$

REMARK. Notice the several ways in which ω affects the sample path of $K_\tau B + S_\tau x$: at time t it takes the value

$$\begin{array}{ll} B(t,\omega) & \text{if } 0 \le t < \tau(\omega) \\ B(\tau(\omega),\omega) + x(t - \tau(\omega)) & \text{if } t \ge \tau(\omega) \end{array}$$

The Brownian filtration

If we regard a Brownian motion $\{B_t : t \in \mathbb{R}^+\}$ as just a Gaussian process with continuus paths and a specific covariance structure (cf. UGMTP §9.3), we need not explicitly mention the filtration. In fact, we could always use the ***natural filtration*** defined by the process itself:

$$\begin{aligned} \mathcal{F}_t^\circ &:= \sigma\{B_s : 0 \le s \le t\} \\ &= \text{sigma-field generated by } K_t B \qquad \text{see Problem [2].} \end{aligned}$$

The process B is adapted to the natural filtration and $\{(B_t, \mathcal{F}_t^\circ) : 0 \le t \le 1\}$ is a Brownian motion in the sense defined by Project 2.

If we augment the filtration by adding the neglible sets to the generating class,

$$\mathcal{F}_t = \sigma\left(\mathcal{F}_t^\circ \cup \mathcal{N}\right),$$

it should be easy for you to check that $\{(B_t, \mathcal{F}_t) : t \in \mathbb{R}^+\}$ is still a Brownian motion.

In fact, B is also a Brownian motion with respect to the standard filtration

$$\widetilde{\mathcal{F}}_t = \mathcal{F}_{t+} = \cap_{s>t}\sigma\left(\mathcal{F}_t^\circ \cup \mathcal{N}\right)$$

Proof.

- Suppose $s < t$ and $F \in \widetilde{\mathcal{F}}_s$. Explain why it is enough to show that
$$\mathbb{P}Ff\left(B_t - B_s\right) = \left(\mathbb{P}F\right)\left(\mathbb{P}f(Z)\right) \qquad \text{where } Z \sim N(0, t-s)$$
for each bounded continuous f.
- Choose a sequence with $t > s_n \downarrow\downarrow s$. Show that $F \in \mathcal{F}_{s_n}$ and
$$\mathbb{P}Ff\left(B_t - B_{s_n}\right) = \left(\mathbb{P}F\right)\left(\mathbb{P}f(Z_n)\right) \qquad \text{where } Z_n \sim N(0, t-s_n).$$
- Pass to the limit. □

BMstd <3> **Corollary.** *The filtration $\{\mathcal{F}_t : t \in \mathbb{R}^+\}$ is standard. That is, $\widetilde{\mathcal{F}}_t = \mathcal{F}_t = \sigma\left(\mathcal{F}_t^\circ \cup \mathcal{N}\right)$ for each t.*

Proof. Suppose $F \in \widetilde{\mathcal{F}}_t$. Then $F \in \mathcal{F}_s$ for each $s > t$. Fix one such s.

- Show there is an $F^\circ \in \mathcal{F}_s^\circ$ for which $F \Delta F^\circ \in \mathcal{N}$.
- Explain why there exists a $\{0,1\}$-valued, $\mathcal{C}$-measurable functional h on $C[0,\infty)$ for which $F^\circ = h(B)$.

- Show that, with probability one,

$$\begin{aligned} F &= \mathbb{P}\left(F \mid \widetilde{\mathcal{F}}_t\right) \\ &= \mathbb{P}\left(h(B) \mid \widetilde{\mathcal{F}}_t\right) \\ &= \mathbb{W}^x h(K_t B + S_t x). \end{aligned}$$

- Explain why $\mathbb{W}^x h(K_t B + S_t x)$ is a $\mathcal{C}$-measurable function of $K_t B$ and hence it is $\mathcal{F}_t^\circ$-measurable.
- Conclude that $F \in \sigma(\mathcal{N} \cup \mathcal{F}_t^\circ) = \mathcal{F}_t$.

□

Problems = some possible examples for your notes

cylinder.Borel [1] One metric for uniform convergence on compacta of function in $C[0,\infty)$ is defined by

$$d(x,y) := \sum_{n\in\mathbb{N}} 2^{-n} \min\left(1, \sup_{0\le t\le n} |x(t) - y(t)|\right)$$

Show that the Borel sigma-field for this metric is exactly the cylinder sigma-field $\mathcal{C}$.

raw [2] Suppose X is a stochastic process with sample paths in $C[0,\infty)$. For each fixed t, define $\mathcal{F}_t^\circ := \sigma\{X_s : 0 \le s \le t\}$. Show that $\mathcal{F}_t^\circ$ is the smallest sigma-field for which the map $\omega \mapsto K_t X(\cdot,\omega)$ is $\mathcal{F}_t^\circ\backslash\mathcal{C}$-measurable.

References

McLeish, D. L. (1974), 'Dependent central limit theorems and invariance principles', *Annals of Probability* **2**, 620–628.

PROJECT 4

Things to explain in your notebook:

(i) How to construct the isometric stochastic integral for a square integrable martingale.

(ii) What advantages are there to considering only predictable integrands?

(iii) Why does it suffice to have the Doléans measure defined only on the predictable sigma-field?

Notation and facts:

- Fixed complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Standard filtration.
- ***R-process*** = adapted process with cadlag sample paths
- ***L-process*** = adapted process with left-continuous sample paths with finite right limits

borrowed from Rogers & Williams (1987)

- $\mathcal{M}^2 = \mathcal{M}^2[0,1]$ = martingales with index set $[0,1]$, cadlag sample paths, and $\mathbb{P}M_1^2 < \infty$ ("square integrable martingales")
- $\mathcal{M}_0^2 = \mathcal{M}_0^2[0,1] = \{M \in \mathcal{M}^2[0,1] : M_0 \equiv 0\}$
- $\mathcal{H}_{\text{simple}}$ = the set of all ***simple processes*** of the form

<1>
$$\sum\nolimits_{i=0}^{N} h_i(\omega)\{t_i < t \le t_{i+1}\}$$

for some grid $0 = t_0 < t_1 < \ldots < t_{N+1} = 1$ and bounded, $\mathcal{F}(t_i)$-measurable random variables h_i. Note that $\mathcal{H}_{\text{simple}}$ is a subset of the set of all L-processes.

REMARK. Some authors call members of $\mathcal{H}_{\text{simple}}$ ***elementary processes***; others reserve that name for the situation where the t_i are replaced by stopping times. Dellacherie & Meyer (1982, §8.1) adopted the opposite convention.

- Abbreviate $\mathbb{P}(\ldots \mid \mathcal{F}_s)$ to $\mathbb{P}_s(\ldots)$.
- Doob's inequality: $\mathbb{P}\sup_{0\le t\le 1} M_t^2 \le 4\mathbb{P}M_1^2$ for $M \in \mathcal{M}^2[0,1]$.

cf. UGMTP Problem 6.9

Increasing processes as measures

Suppose $M \in \mathcal{M}^2$ is such that there exists an R-process A with increasing sample paths such that the process $N_t := M_t^2 - A_t$ is a martingale. Without loss of generality, $M_0 = A_0 = 0$. For example, for Brownian motion, $A_t \equiv t$.

REMARK. The existence of such an A for each M in $\mathcal{M}^2[0,1]$ will follow later from properties of stochastic integrals. See the discussion of quadratic variation.

Identify $A(\cdot, \omega)$ with a measure μ_ω on $\mathcal{B}(0,1]$ for which

$$\mu_\omega(0,t] = A(t,\omega) \qquad \text{for } 0 < t \le 1$$

Construct a measure μ on $\mathcal{B}(0,1] \otimes \mathcal{F}$ by

$$\mu g(t,\omega) = \mathbb{P}^\omega \mu_\omega^t g(t,\omega) \qquad \text{for which } g?$$

Notice that $\mu(0,1] \times \Omega = \mathbb{P}A_1 < \infty$.

- For Brownian motion, show that $\mu = \mathfrak{m} \otimes \mathbb{P}$ with $\mathfrak{m}$ = Lebesgue measure on $\mathcal{B}(0,1]$.
- For fixed $0 \le a < b \le 1$, define $\Delta N = N_b - N_a$, $\Delta M = M_b - M_a$, and $\Delta A = A_b - A_a$. Show that

$$0 = \mathbb{P}_a \Delta N = \mathbb{P}_a\left((\Delta M)^2 - \Delta A\right) \qquad \text{almost surely.}$$

- At least for each bounded, $\mathcal{F}_a$-measurable random variable h, deduce that

$$\mathbb{P}h(\omega)(\Delta M)^2 = \mathbb{P}h(\omega)\Delta A = \mathbb{P}^{\omega}\left(h(\omega)\mu_{\omega}^{t}\{a < t \le b\}\right)$$

<2> $$= \mu h(\omega)\{a < t \le b\}$$

Stochastic integral for simple processes

Suppose H is a simple process, as in <1>, and $M \in \mathcal{M}^2$. The stochastic integral is defined by

<3> $$\int_{(0,1]} H\,dM := \sum\nolimits_{i=0}^{N} h_i(\omega)\left(M(t_{i+1},\omega) - M(t_i,\omega)\right).$$

Remark. Here I follow Rogers & Williams (1987, page 2) in excluding the lower endpoint from the range of integration. Dellacherie & Meyer (1982, §8.1) added an extra contribution from a possible jump in M at 0. With the $(0,1]$ interpretation, the definition depends only on the increments of M; with no loss of generality, we may therefore assume $M_0 \equiv 0$.

A similar awkwardness arises in defining $\int_0^t H\,dM$ if M has a jump at t. The notation does not distinguish between the integral over $(0,t)$ and the integral over $(0,t]$. I will use instead the Strasbourg notation $H \bullet M_1$ for $\int_{(0,1]} H\,dM$, with H multiplied by an explicit indicator function to modify the range of integration. For example, $\int_0^t H\,dM$ is obtained from <3> by substituting $H(s,\omega)\{0 < s \le t\}$ for H. Thus,

Check

<4> $$H \bullet M_t := \sum\nolimits_{i=0}^{N} h_i(\omega)\left(M(t \wedge t_{i+1},\omega) - M(t \wedge t_i,\omega)\right).$$

- You should check that $\mathbb{P}_t H \bullet M_1 = H \bullet M_t$ almost surely, so that $H \bullet M$ is a martingale (with cadlag paths).

<5> **Lemma.** *$\mathbb{P}\left(H \bullet M_1\right)^2 = \mu H^2$ for each $H \in \mathcal{H}_{\text{simple}}$,*

Proof. Expand the left-hand side of the asserted inequality as

$$\sum\nolimits_i \mathbb{P}h_i^2(\Delta_i M)^2 + 2\sum\nolimits_{i<j} \mathbb{P}h_i h_j \Delta_i M \Delta_j M \qquad \text{where } \Delta_i M = M(t_{i+1} - M(t_i).$$

Use the fact that $\mathbb{P}(\Delta_j M \mid \mathcal{F}(t_{j-1})) = 0$ to kill all the cross-product terms. Use equality <2> to simplify the other contributions to

$$\mu^{s,\omega}\sum\nolimits_i h_i(\omega)^2\{t_i < s \le t_{i+1}\} = \mu H^2$$

□

Extension by isometry

Think of $\mathcal{H}_{\text{simple}}$ as a subspace of $\mathcal{L}^2 = \mathcal{L}^2((0,1]\times\Omega, \mathcal{B}(0,1]\otimes\mathcal{F}_1, \mu)$. Then Lemma <5> shows that $H \mapsto H \bullet M_1$ is an isometry from a subspace of $\mathcal{L}^2$ to $\mathcal{L}^2(\Omega,\mathcal{F}_1,\mathbb{P})$. It extends to an isometry from $\overline{\mathcal{H}}_{\text{simple}}$, the $\mathcal{L}^2(\mu)$ closure of $\mathcal{H}_{\text{simple}}$ in $\mathcal{L}^2$, into $\mathcal{L}^2(\Omega,\mathcal{F}_1,\mathbb{P})$. The stochastic integral $H \bullet M_t$ is then taken to be a cadlag version of the martingale $\mathbb{P}_t H \bullet M_1$. In short, there is a linear map $H \mapsto H \bullet M$ from $\overline{\mathcal{H}}_{\text{simple}}$ to $\mathcal{M}_0^2$ for which, by Doob's inequality,

<6> $$\mathbb{P}\sup_{0\le t\le 1}|G \bullet M_t - H \bullet M_t|^2 \le 4\mathbb{P}|H \bullet M_1 - G \bullet M_1|^2 = \mu|G - H|^2$$

It is uniquely determined by the property, for all $a < b$ and $F \in \mathcal{F}_a$,

$$H \bullet M_1 = F\left(M_b - M_a\right) \qquad \text{if } H(t,\omega) = \{\omega \in F\}\{a < t \le b\}.$$

<7> **Example.** Let τ be a stopping time taking values in $[0, 1]$. Define the stochastic interval

$$((0, \tau]] := \{(t, \omega) \in (0, 1] \times \Omega : 0 < t \le \tau(\omega)\}$$

Let τ_n be the stopping time obtained by rounding τ up to the next integer multiple of 2^{-n}:

$$\tau_n(\omega) = \sum\nolimits_{i=1}^{2^n} t_i\{t_{i-1} < \tau(\omega) \le t_i\} \qquad \text{where } t_i = i/2^n.$$

- Show that
$$((0, \tau_n]] = \sum\nolimits_{i=1}^{2^n} \{t_{i-1} < t \le t_i\}\{\tau(\omega) > t_{i-1}\} \in \mathcal{H}_{\text{simple}}$$
and that $\mu\left(((0, \tau_n]] - ((0, \tau]]\right)^2 \to 0$.
- Conclude that $((0, \tau]] \bullet M_t = M_{t\wedge\tau}$. □

Predictable integrands

How large is $\overline{\mathcal{H}}_{\text{simple}}$? For Brownian motion, it s traditional to show (Chung & Williams 1990, Theorem 3.7) that $\overline{\mathcal{H}}_{\text{simple}}$ contains at least all the $\mathcal{B}(0, 1] \times \mathcal{F}_1$-measurable, adapted processes that are square integrable for $\mathfrak{m} \times \mathbb{P}$. For other martingales, it is cleaner to work with a slightly smaller class of integrands.

<8> **Definition.** *The predictable sigma-field $\mathcal{P}$ is defined as the sigma-field on $(0, 1] \times \Omega$ generated by the set of all L-processes. The space $\mathcal{H}^2(\mu)$ is defined as the set of all $\mathcal{P}$-measurable processes H on $(0, 1] \times \Omega$ for which $\mu H^2 < \infty$.*

Notice that $\mathcal{H}_{\text{simple}} \subseteq \mathcal{H}^2(\mu)$ for the μ corresponding to each M in $\mathcal{M}_0^2$. In fact, a generating class argument shows that $\mathcal{H}^2(\mu)$ is the closure of $\mathcal{H}_{\text{simple}}$ in the space $\mathcal{L}^2(\,(0, 1] \times \Omega, \mathcal{P}, \mu)$:

- Suppose H is a bounded, L-process. Define
$$H_n(t, \omega) := \sum\nolimits_{i=1}^{2^n} H(t_{i-1}, \omega)\{t_{i-1} < t \le t_i\} \qquad \text{where } t_i = i/2^n$$
Show that $H_n \in \mathcal{H}_{\text{simple}}$ and that $H_n(t, \omega) \to H(t, \omega)$ for all (t, ω) and hence that $\mu\left(H_n - H\right)^2 \to 0$. Deduce that $H \in \overline{\mathcal{H}}_{\text{simple}}$.
- Invoke a generating class argument (such as the one given in the extract *generating-class-fns.pdf* from UGMTP) to deduce that $\overline{\mathcal{H}}_{\text{simple}}$ contains all bounded, $\mathcal{P}$-measurable processes.
- Then what?

The Doléans measure

If we intend only to extend the stochastic integral to predictable integrands, we do not need the measure μ that corresponds to the increasing process A to be defined on $\mathcal{B}(0, 1] \otimes \mathcal{F}_1$: we only need it defined on $\mathcal{P}$. In fact, it is a much easier task to construct an appropriate μ on $\mathcal{P}$ directly from the submartingale $\{M_t^2 : 0 \le t \le 1\}$ without even assuming the existence of A. The measure μ is called the ***Doléans measure*** for the submartingale M^2. See the handout *Doleans.pdf* for a construction.

Moreover, there is another procedure (the dual predictable projection) for extending the Doléans measure to a "predictable measure" on $\mathcal{B}(0, 1] \otimes \mathcal{F}_1$. A disintegration of this new measure then defines the process A. I'll prepare a handout describing the method.

Problems

[1] Show that the predictable sigma-field $\mathcal{P}$ on $(0,1] \times \Omega$ is generated by each of the following sets of processes:

(i) all sets $(a,b] \times F$ with $F \in \mathcal{F}_a$ and $0 \le a < b \le 1$

(ii) $\mathcal{H}_{\text{simple}}$

(iii) the set $\mathbb{C}$ of all adapted processes with continuous sample paths

(iv) all stochastic intervals $((0,\tau]]$ for stopping times τ taking values in $[0,1]$

(v) all sets $\{(t,\omega) \in (0,1] \times \Omega : X(t,\omega) = 0\}$, with $X \in \mathbb{C}$

References

Chung, K. L. & Williams, R. J. (1990), *Introduction to Stochastic Integration*, Birkhäuser, Boston.

Dellacherie, C. & Meyer, P. A. (1982), *Probabilities and Potential B: Theory of Martingales*, North-Holland, Amsterdam.

Rogers, L. C. G. & Williams, D. (1987), *Diffusions, Markov Processes, and Martingales: Itô Calculus*, Vol. 2, Wiley.

Project 5

This week I would like you to consolidate your understanding of the material from the last two weeks by working through some problems.

- Read the handout *Doleans.pdf,* at least up to Theorem 5. Try to explain the assertions flagged by the symbol $\Rrightarrow$. Try to solve Problem [2].

<1> **Definition.** *Suppose X is a process and τ is a stopping time. Define the **stopped process** $X_{\wedge\tau}$ to be the process for which $X_{\wedge\tau}(t,\omega) = X(\tau(\omega) \wedge t, \omega)$.*

nonstandard notation

Problems

[1] Suppose $M \in \mathcal{M}^2[0,1]$ has continuous sample paths.

(i) For each H in $\mathcal{H}_{\text{simple}}$, show that $H \bullet M$ has continuous sample paths.

(ii) Suppose $\{H_n : n \in \mathbb{N}\} \subseteq \mathcal{H}_{\text{simple}}$ and $\mu|H_n - H|^2 \to 0$. Use Doob's maximal inequality to show that there exists a subsequence $\mathbb{N}_1$ along which

$$\sum_{n\in\mathbb{N}_1} \mathbb{P} \sup_{0\le t\le 1} |H_n \bullet M_t - H \bullet M_t| < \infty$$

(iii) Deduce that there is a version of $H \bullet M$ with continuous sample paths.

[2] Suppose $\{\tau_n : n \in \mathbb{N}\}$ is a sequence of $[0,1]$-valued stopping times for which $\tau_n \le \tau$ and $\tau_n(\omega) \uparrow\uparrow \tau(\omega)$ at each ω for which $\tau(\omega) > 0$. Prove that $[[\tau, 1]] \in \mathcal{P}$.

Remark. The sequence $\{\tau_n\}$ is said to ***foretell*** τ. Existence of a foretelling sequence was originally used to define the concept of a ***predictable stopping time***. The modern definition requires $[[\tau, 1]] \in \mathcal{P}$. In fact, the two definitions are almost equivalent, as you will see from a later handout.

[3] Suppose $M \in \mathcal{M}_0^2[0,1]$ and $H \in \mathcal{H}^2(\mu)$, where μ is the Doléans measure defined by the submartingale M_t^2. Let τ be a $[0,1]$-valued stopping time. Let X denote the martingale $H \bullet M$.

(i) Define $N = M_{\wedge\tau}$. Show that $N \in \mathcal{M}_0^2[0,1]$.

(ii) Show that, with probability one,

$$X_{t\wedge\tau} = \Big(\big(H((0,\tau]]\big) \bullet M \Big)_t = H \bullet N_t \qquad \text{for } 0 \le t \le 1.$$

Hint: Consider first the case where $H \in \mathcal{H}_{\text{simple}}$ and τ takes values only in a finite subset of $[0,1]$. Extend to general τ by rounding up to integer multiples of 2^{-n}.

(iii) Show that the Doléans measure ν for the submartingale $(H \bullet M)_t^2$ has density H^2 with respect to μ. Hint: Remember that the Doléans measure is uniquely determined by the values it gives to the stochastic intervals $((0,\tau]]$.

(iv) Suppose $K \in \mathcal{H}^2(\nu)$. Show that $KH \in \mathcal{H}^2(\mu)$ and $K \bullet (H \bullet M) = (HK) \bullet M$.

[4] Suppose $\mu = \mathfrak{m} \otimes \mathbb{P}$, defined on $\mathcal{B}(0,1] \otimes \mathcal{F}_1$. Let $\{X_t : 0 \le t \le 1\}$ be progressively measurable.

(i) Suppose X is bounded, that is, $\sup_{t,\omega} |X(t,\omega)| < \infty$. Define

$$H_n(t,\omega) := n \int_{t-n^{-1}}^{t} X(s,\omega)\, ds$$

(How should you understand the definition when $t < n^{-1}$?) Show that H_n is predictable and that $\int_0^1 |H_n(t,\omega) - X(t,\omega)|^2\,dt \to 0$ for each ω.

(ii) Deduce that $\mu|H_n - X|^2 \to 0$.

(iii) Deduce that $X \in \overline{\mathcal{H}}_{\text{simple}}$, the closure in $\mathcal{L}^2(\mathcal{B}(0,1]\otimes\mathcal{F}_1,\mu)$, if $\mu X^2 < \infty$.

[5] Suppose $M \in \mathcal{M}_0^2[0,1]$, with μ the Doléans measure defined by the submartingale M_t^2. Suppose $\psi : \mathcal{H}^2(\mu) \to \mathcal{M}_0^2[0,1]$ is a linear map, in the sense that $\psi(\alpha H + \beta K)_t = \alpha\psi(H)_t + \beta\psi(K)_t$ almost surely, for each $t \in [0,1]$ and constants $\alpha, \beta \in \mathbb{R}$. Suppose that

(a) $\psi(((0,\tau]])_t = M_{t\wedge\tau}$ almost surely, for each stopping time τ.

(b) if $1 \ge |H_n| \to 0$ pointwise then $\psi(H_n)_t \to 0$ in probability, for each fixed t.

Show that these properties characterize the stochastic integral, in the following senses.

(i) Show that $\psi(H)_t = H \bullet M_t$ almost surely, for each t. Hint: Consider the collection of all bounded, nonnegative, predictable processes H for which $\psi(H)_t = H \bullet M_t$ almost surely, for every t. Use a generating class argument.

(ii) If, in addition, $\psi(H_n)_t \to 0$ in probability whenever $\mu H_n^2 \to 0$, show that the conclusion from part (i) also holds for every H in $\mathcal{H}_2(\mu)$.

Project 6

This week there are many small details that might occcupy your attention. I would be satisfied if you concentrated on some of the more important points. Things to explain in your notebook:

(i) Why is the theory for $\mathcal{M}_0^2(\mathbb{R}^+)$ almost the same as the theory for $\mathcal{M}_0^2[0,1]$?

(ii) Why do we get sigma-finite Doléans measures for the submartingales corresponding to $\text{loc}\mathcal{M}_0^2(\mathbb{R}^+)$ processes?

(iii) Why can $H \bullet M$ be built up pathwise from isometric stochastic integrals when $H \in \text{loc}\mathcal{H}_{\text{Bdd}}$ and $M \in \text{loc}\mathcal{M}_0^2(\mathbb{R}^+)$?

(iv) Why do we need to replace $\mathcal{L}^2(\mathbb{P})$ convergence by convergence in probability after localizing?

Square-integrable martingales indexed by $\mathbb{R}^+$

- Define $\mathcal{M}^2(\mathbb{R}^+)$ as the set of all square-integrable martingales, that is, cadlag martingales $\{M_t : t \in \mathbb{R}^+\}$ for which $\sup_t \mathbb{P}M_t^2 < \infty$.
- Define $\mathcal{F}_\infty = \sigma\left(\cup_{t\in\mathbb{R}^+} \mathcal{F}_t\right)$. If $M \in \mathcal{M}^2(\mathbb{R}^+)$ then there exists an $M_\infty \in \mathcal{L}^2(\Omega, \mathcal{F}_\infty, \mathbb{P})$ such that $M_t \to M_\infty$ almost surely and $\mathbb{P}|M_t - M_\infty|^2 \to 0$ as $t \to \infty$. Moreover, $M_t = \mathbb{P}(M_\infty \mid \mathcal{F}_t)$ almost surely.
- Note that $\{(M_t, \mathcal{F}_t) : 0 \le t \le \infty\}$ is also a martingale. Suppose ψ is a one-to-one map from $[0,1]$ onto $[0,\infty]$, such as $\psi(s) = s/(1-s)$. Define $\mathcal{G}_s := \mathcal{F}(\psi(s))$ and $N_s = M(\psi(s))$. Then $\{(N_s, \mathcal{G}_s) : 0 \le s \le 1\}$ belongs to $\mathcal{M}^2[0,1]$. All the theory for the isometric stochastic integrals with respect to $\mathcal{M}^2[0,1]$ processes carries over to analogous theory for $\mathcal{M}^2(\mathbb{R}^+)$.
- Note a subtle difference: For $\mathcal{M}^2(\mathbb{R}^+)$ we have left continuity of sample paths at ∞, by construction of M_∞. For $\mathcal{M}^2[0,1]$ we did not require left continuity at 1. Also we did not require that $\mathcal{F}_1 = \sigma\left(\cup_{t<1} \mathcal{F}_t\right)$. A better analogy would allow $\mathcal{F}_\infty$ to be larger than $\sigma\left(\cup_{t\in\mathbb{R}^+} \mathcal{F}_t\right)$ and would allow M to have a jump at ∞.

Localization

<1> **Definition.** *Suppose X is a process and τ is a stopping time. Define the* ***stopped process*** $X_{\wedge\tau}$ *to be the process for which $X_{\wedge\tau}(t,\omega) = X(\tau(\omega) \wedge t, \omega)$.*

nonstandard notation

<2> **Definition.** *Suppose $\mathcal{W}$ is a set of processes (indexed by $\mathbb{R}^+$) that is stable under stopping, $W \mapsto W_{\wedge\tau}$. Say that a process X is locally in $\mathcal{W}$ if there exists a sequence of stopping times $\{\tau_k\}$ with $\tau_k \uparrow \infty$ and $X_{\wedge\tau_k} \in \mathcal{W}$ for each k. Call $\{\tau_k\}$ a* ***$\mathcal{W}$-localizing sequence*** *for X. Write $\text{loc}\mathcal{W}$ for the set of all processes that are locally in $\mathcal{W}$.*

or reducing sequence

Remark. Notice that if $\{\tau_k\}$ is a $\mathcal{W}$-localizing sequence for X then so is $\{k \wedge \tau_k\}$. Thus we can always require each τ_k in a localizing sequence to be a bounded stopping time.

Predictable sigma-field

The predictable sigma-field $\mathcal{P}$ on $(0,\infty) \times \Omega$ is again defined as the sigma-field generated by all L-processes.

Remark. For $(0,1] \times \Omega$ the predictable sigma-field contains some subsets of $\{1\} \times \Omega$. For $(0,\infty) \times \Omega$, subsets of $\{\infty\} \times \Omega$ are not in $\mathcal{P}$. Maybe it would be better to define $\mathcal{P}$ on $(0,\infty] \times \Omega$.

Stochastic intervals

For stopping times σ and τ taking values in $\mathbb{R}^+ \cup \{\infty\}$ define

$$((\sigma, \tau]] := \{(t, \omega) \in \mathbb{R}^+ \times \Omega : \sigma(\omega) < t \leq \tau(\omega)\},$$

and so on. Note well that the stochastic interval is a subset of $\mathbb{R}^+ \times \Omega$. Points (t, ω) with $t = \infty$ are not included, even at ω for which $\tau(\omega) = \infty$. In particular, for $\sigma \equiv 0$ and $\tau \equiv \infty$ we get

$$((0, \infty]] = \mathbb{R}^+ \times \Omega.$$

Don't be misled by the "$\infty]]$" into assuming that $\{\infty\} \times \Omega$ is included.

> REMARK. The convention that ∞ is excluded makes possible some neat arguments, even though it spoils the analogy with stochastic subintervals of $(0, 1] \times \Omega$. Although sorely tempted to buck tradition, I decided to stick with established usage for fear of unwanted exceptions to established theorems.

- Write $\mathcal{T}$ for the set of all $[0, \infty]$-valued stopping times. Is it true that $\mathcal{P}$ is generated by the set of all stochastic intervals $((0, \tau]]$ for $\tau \in \mathcal{T}$?
- If $M \in \mathcal{M}_0^2(\mathbb{R}^+)$ explain why there exists a finite, countably-additive measure on $\mathcal{P}$ (the Doléans measure for the submartingale M^2) for which

  $$\mu(a, b] \times F = \mathbb{P}F(M_b - M_a)^2 \qquad \text{for } F \in \mathcal{F}_a \text{, and } 0 \leq a < b < \infty.$$

  Could we also allow $b = \infty$? Is it still true that

  $$\mu((0, \tau]] = \mathbb{P}M_\tau^2 \qquad \text{for each } \tau \in \mathcal{T}?$$

  How should the last equality be interpreted when $\{\omega : \tau(\omega) = \infty\} \neq \emptyset$?

Locally square-integrable martingales

- Consider first the case of a process M for which there exists a stopping time σ such that $N := M_{\wedge\sigma} \in \mathcal{M}^2(\mathbb{R}^+)$. Let μ be the Doléans measure on $\mathcal{P}$ for the square-integrable submartingale N^2.
  - (i) Is it true that $N_\infty = M_\sigma$? What would this equality be asserting about those ω at which $\sigma(\omega) = \infty$?
  - (ii) Show that $\mu((0, \infty]] = \sup_t \mathbb{P}M_{t\wedge\sigma}^2 = \mathbb{P}N_\infty^2$.
  - (iii) Show that $\mu((t \wedge \sigma, \infty]] = \mathbb{P}(N_\infty - M_{t\wedge\sigma})^2 \to 0$ as $t \to \infty$.
  - (iv) Conclude that μ is a finite measure that concentrates all its mass on the stochastic interval $((0, \sigma]]$.
- Now suppose $M \in \text{loc}\mathcal{M}_0^2(\mathbb{R}^+)$, with localizing sequence $\{\tau_k : k \in \mathbb{N}\}$. Write μ_k for the Doléans measure of the submartingale $M_{\wedge\tau_k}^2$.
  - (i) Show that μ_k is a finite measure concentrating on $((0, \tau_k]]$ and that the restriction of μ_{k+1} to $((0, \tau_k]]$ equals μ_k.
  - (ii) Define μ on $\mathcal{P}$ by $\mu H := \sup_{k\in\mathbb{N}} \mu_k H$. Show that μ is a sigma-finite, countably-additive measure for which

    $$\mu((0, \tau]] = \sup_k \mathbb{P}M_{\tau\wedge\tau_k}^2 \qquad \text{for all } \tau \in \mathcal{T}.$$

  - (iii) Suppose $\{\sigma_k : k \in \mathbb{N}\}$ is another localizing sequence for M. Show that

    $$\mu((0, \tau]] = \sup_k \mathbb{P}M_{\tau\wedge\sigma_k}^2 \qquad \text{for all } \tau \in \mathcal{T}.$$

    That is, show that μ does not depend on the choice of localizing sequence for M.

Locally bounded predictable processes

Write $\mathcal{H}_{\text{Bdd}}$ for the set of all bounded, $\mathcal{P}$-measurable processes.

- Show that every L-process X with $\sup_\omega |X_0(\omega)| < \infty$ belongs to $\text{loc}\mathcal{H}_{\text{Bdd}}$. Hint: Consider $\tau_k(\omega) := \inf\{t \in \mathbb{R}^+ : |X_t(\omega)| \geq k\}$.

  REMARK. Does an L-process have time set $[0, \infty)$ or $(0, \infty)$? Perhaps the assertion would be better expressed as: the restriction of X to $(0, \infty) \times \Omega$ belongs to $\text{loc}\mathcal{H}_{\text{Bdd}}$. In that case, the assumption about X_0 is superfluous. D&M have some delicate conventions and definitions for handling contributions from $\{0\} \times \Omega$.

- (Much harder) Is the previous assertion still true if we replace L-processes by $\mathcal{P}$-measurable processes? What if we also require each sample path to be cadlag?

  REMARK. A complete resolution of this question requires some facts about predictable stopping times and predictable cross-sections. Compare with Métivier (1982, Section 6).

Localization of the isometric stochastic integral

The new stochastic integral will be defined indirectly by a sequence of isometries. The continuity properties of $H \bullet M$ will be expressed not via $\mathcal{L}^2$ bounds but by means of the concept of ***uniform convergence in probability on compact intervals***. For a sequence of processes $\{Z_n\}$, write $Z_n \xrightarrow{ucpc} Z$ to mean that $\sup_{0 \leq s \leq t} |Z_n(s, \omega) - Z(s, \omega)| \to 0$ in probability, for each t in $\mathbb{R}^+$.

nonstandard notation

<3> **Theorem.** *Suppose $M \in \text{loc}\mathcal{M}_0^2(\mathbb{R}^+)$. There exists a linear map $H \mapsto H \bullet M$ from $\text{loc}\mathcal{H}_{\text{Bdd}}$ into $\text{loc}\mathcal{M}_0^2(\mathbb{R}^+)$ with the following properties.*

(i) $((0, \tau]] \bullet M_t = M_{t \wedge \tau}$ *for all* $\tau \in \mathcal{T}$.

(ii) $(H \bullet M)_{t \wedge \tau} = \big(H((0, \tau]]\big) \bullet M_t = \big(H \bullet M_{\wedge \tau}\big)_t$, *for all* $H \in \text{loc}\mathcal{H}_{\text{Bdd}}$ *and all* $\tau \in \mathcal{T}$.

(iii) If M has continuous sample paths then so does $H \bullet M$.

*(iv) Suppose $\{H^{(n)} : n \in \mathbb{N}\} \subseteq \text{loc}\mathcal{H}_{\text{Bdd}}$ and $H^{(n)}(t, \omega) \to 0$ for each (t, ω). Suppose that the sequence is **locally uniformly bounded**: there exist stopping times with $\tau_k \uparrow \infty$ and finite constants C_k such that $|H^{(n)}_{\wedge \tau_k}| \leq C_k$ for each n and each k. Then $H^{(n)} \bullet M \xrightarrow{ucpc} 0$.*

(v) $K \bullet (H \bullet M) = (KH) \bullet M$ *for* $K, H \in \text{loc}\mathcal{H}_{\text{Bdd}}$.

Sketch of a proof. Suppose M has localizing sequence $\{\tau_k : k \in \mathbb{N}\}$ and $H \in \text{loc}\mathcal{H}_{\text{Bdd}}$ has localizing sequence $\{\sigma_k : k \in \mathbb{N}\}$.

- Why is there no loss of generality in assuming that $\sigma_k = \tau_k$ for every k?
- Write $M^{(k)}$ for $M_{\wedge \tau_k}$. Define $X^{(k)}$ to be the square integrable martingale
$$X^{(k)} = H_{\wedge \tau_k} \bullet M^{(k)} = \big(H((0, \tau_k]]\big) \bullet M^{(k)}.$$
  Why are the two integrals the same, up to some sort of almost sure equivalence?
- Show that $X^{(k)}(t, \omega) = X^{(k)}(t \wedge \tau_k(\omega), \omega)$ for all $t \in \mathbb{R}^+$. That is, show that the sample paths are constant for $t \geq \tau_k(\omega)$. Do we need some sort of almost sure qualification here?
- Show that, on a set of ω with probability one,
$$X^{(k+1)}(t \wedge \tau_k(\omega), \omega) = X^{(k)}(t \wedge \tau_k(\omega), \omega) \qquad \text{for all } t \in \mathbb{R}^+.$$
- Show that there is an R-process X for which, on a set of ω with probability one,
$$X(t \wedge \tau_k(\omega), \omega) = X^{(k)}(t \wedge \tau_k(\omega), \omega) \qquad \text{for all } t \in \mathbb{R}^+, \text{ all } k.$$

- Show that $X \in \text{loc}\mathcal{M}_0^2(\mathbb{R}^+)$, with localizing sequence $\{\tau_k : k \in \mathbb{N}\}$.
- Define $H \bullet M := X$.
- In order to establish linearity of $H \mapsto H \bullet M$, we need to show that the definition does not depend on the particular choice of the localizing sequence. (If we can use a single localizing sequence for two different H processes then linearity for the approximating $X^{(k)}$ processes will transfer to the X process.)
- For assertion (iv), we may also assume that $\{\tau_k\}$ localizes M to $\mathcal{M}_0^2(\mathbb{R}^+)$. Write μ_k for the Doléans measure of the submartingale $(M^{(k)})^2$. Then, for each fixed k, we have

$$\begin{aligned}\mathbb{P}\sup_{s\le t}\left(H^{(n)} \bullet M\right)^2_{s\wedge\tau_k} &= \mathbb{P}\sup_{s\le t}\left(H^{(n)}((0,\tau_k]] \bullet M^{(k)}\right)^2_s && \text{by construction}\\ &\le 4\mathbb{P}\left(H^{(n)}((0,\tau_k]] \bullet M^{(k)}\right)^2_t && \text{by Doob's inequality}\\ &= 4\mu_k\left((H^{(n)})^2((0,\tau_k\wedge t]]\right)\\ &\to 0 \qquad \text{as } n\to\infty, \text{ by Dominated Convergenece.}\end{aligned}$$

When $\tau_k > t$, which happens with probability tending to one, the processes $H^{(n)} \bullet M_{s\wedge\tau_k}$ and $H^{(n)} \bullet M_s$ coincide for all $s \le t$. The uniform convergence in probability follows.

□

Characterization of the stochastic integral

<4> **Theorem.** *Suppose $M \in \mathcal{M}_0^2(\mathbb{R}^+)$. Suppose also that $\psi : \text{loc}\mathcal{H}_{\text{Bdd}} \to \mathcal{M}_0^2(\mathbb{R}^+)$ is a linear map (in the sense of almost sure equivalence) for which*

better just to state equality for bounded τ?

(i) $((0,\tau]] \bullet M_t = M_{\tau\wedge t}$ almost surely, for each $t \in \mathbb{R}^+$ and $\tau \in \mathcal{T}$

(ii) If $\{H^{(n)} : n \in \mathbb{N}\} \subseteq \text{loc}\mathcal{H}_{\text{Bdd}}$ is locally uniformly bounded and $H^{(n)}(t,\omega) \to 0$ for each (t,ω) then $\psi(H^{(n)})_t \to 0$ in probability for each t.

Then $\psi(H)_t = H \bullet M_t$ almost surely for each $t \in \mathbb{R}^+$ and each $H \in \text{loc}\mathcal{H}_{\text{Bdd}}$.

Remark. The assertion of the Theorem can also be written: there exists a set Ω_0 with $\Omega_0^c \in \mathcal{N}$ such that

$$\psi(H)(t,\omega) = H \bullet M(t,\omega) \qquad \text{for every } t \text{ if } \omega \in \Omega_0$$

Cadlag sample paths allow us to deduce equality of whole paths from equality ot a countable dense set of times.

References

Métivier, M. (1982), *Semimartingales: A Course on Stochastic Processes*, De Gruyter, Berlin.

Project 7

This week you should concentrate on understanding Theorem <4>, which states the basic properties of integrals with respect to semimartingales. The facts about finite variation are mostly for background information; you could safely regard an $\mathcal{FV}$-process to be defined as a difference of two increasing R-processes.

The facts about quadratic variation process will be used in the next Project to establsh the Itô formula. You might prefer to postpone your careful study of $[X, Y]$ until that Project.

I do not expect you to work every Problem.

1. Cadlag functions of bounded variation

Suppose f is a real function defined on $\mathbb{R}^+$. For each finite grid

$$\mathbb{G}: \quad a = t_0 < t_1 < \ldots < t_N = b$$

on $[a, b]$ define the variation of f over the grid to be

$$V_f(\mathbb{G}, [a, b]) := \sum\nolimits_{i=1}^{N} |f(t_i) - f(t_{i-1})|$$

Say that f is of ***bounded variation*** on the interval $[a, b]$ if there exists a finite constant $V_f[a, b]$ for which

$$\sup\nolimits_{\mathbb{G}} V_f(\mathbb{G}, [a, b]) \le V_f[a, b]$$

where the supremum is taken over the set of all finite grids $\mathbb{G}$ on $[a, b]$. Say that f is of ***finite variation*** if it is of bounded variation on each bounded interal $[0, b]$.

Problems [1] and [2] establish the following facts about finite variation. Every difference $f = f_1 - f_2$ of two increasing functions is of finite variation. Conversely, if f is of finite variation then the functions $t \mapsto V_f[0, t]$ and $t \mapsto V_f[0, t] - f(t)$ are both increasing and

$$f(t) = V_f[0, t] - \big(V_f[0, t] - f(t)\big),$$

a difference of two increasing functions. Moreover, if f is cadlag then $V_f[0, t]$ is also cadlag.

2. Processes of finite variation as random (signed) measures

Let $\{L_t : t \in \mathbb{R}^+\}$ be an R-process with increasing sample paths, adapted to a standard filtration $\{\mathcal{F}_t : t \in \mathbb{R}^+\}$. For each ω, the function $L(\cdot, \omega)$ defines a measure on $\mathcal{B}(\mathbb{R}^+)$,

Notation: $L_t(\omega) = L(t, \omega)$.

$$\lambda_\omega[0, t] = L_t(\omega) \qquad \text{for } t \in \mathbb{R}^+.$$

The family $\Lambda = \{\lambda_\omega : \omega \in \Omega\}$ may be thought of as a ***random measure***, that is, a map from Ω into the space of (sigma-finite) measures on $\mathcal{B}(\mathbb{R}^+)$.

Notice that $\lambda_\omega\{0\} = L_0(\omega)$, an atom at the origin, which can be awkward. It will be convenient if $L_0 \equiv 0$, ensuring that λ_ω concentrates on $(0, \infty)$.

<1> **Definition.** *Write $\mathcal{FV}$, or $\mathcal{FV}(\mathbb{R}^+)$ if there is any ambiguity about the time set, for the set of all R-processes with sample paths that are of finite variation on $\mathbb{R}^+$. Write $\mathcal{FV}_0$ for the subset of $\mathcal{FV}$-processes, A, with $A_0 \equiv 0$.*

- Show that $\mathcal{FV}$ could also be defined as the set of processes expressible as a difference $A(\cdot,\omega) = L'(\cdot,\omega) - L''(\cdot,\omega)$ of two increasing R-processes.

The stochastic integral with respect to A will be defined as a difference of stochastic integrals with respect to L' and L''. Questions of uniqueness—lack of dependence on the choice of the two increasing processes—will be subsumed in the the uniqueness assertion for semimartingales.

The case where L is an increasing R-process with $L_0 \equiv 0$ will bring out the main ideas. I will leave to you the task of extending the results to a difference of two such processes. Define the stochastic integral with respect to L pathwise,

$$H \bullet L_t := \lambda_\omega^s \left(\{0 < s \le t\} H(s,\omega)\right).$$

This integral is well defined if $H \in \text{loc}\mathcal{H}_{\text{Bdd}}$.

Indeed, suppose $H \in \text{loc}\mathcal{H}_{\text{Bdd}}$. There exist stopping times $\tau_k \uparrow \infty$ and finite constants C_k for which $|H_{\wedge\tau_k}| \le C_k$. For each fixed ω, the function $s \mapsto H(s,\omega)$ is measurable (by Fubini, because predictable implies progressively measurable). Also $\sup_{s\le t} |H(s,\omega)| \le C_k$ when $t \le \tau_k(\omega)$. The function $H(\cdot,\omega)$ is integrable with respect to λ_ω on each bounded interval. Moreover, we have a simple bound for the contributions from the positive and negative parts of H to the stochastic integral:

$$0 \le H^\pm \bullet L_{t\wedge\tau_k} = \lambda_\omega^s \left(\{0 \le s \le t \wedge \tau_k\} H^\pm(s,\omega)\right) \le C_k L(t,\omega).$$

That is, $H^\pm \bullet L_t \le C_k L(t,\omega)$ when $t \le \tau_k(\omega)$.

- Show that the sample paths of $H^\pm \bullet L$ are cadlag and adapted. Deduce that $H \bullet L \in \mathcal{FV}_0$.

You should now be able to prove the following result by using standard facts about measures.

<2> **Theorem.** *Suppose $A \in \mathcal{FV}_0$. There is a map $H \mapsto H \bullet A$ from $\text{loc}\mathcal{H}_{\text{Bdd}}$ to $\mathcal{FV}_0$ that is linear (in the almost sure sense?) for which:*

(i) $((0,\tau]] \bullet A_t = A_{t\wedge\tau}$ *for each $\tau \in \mathcal{T}$ and $t \in \mathbb{R}^+$.*

(ii) $(H \bullet A)_{t\wedge\tau} = (H((0,\tau]]) \bullet A_t = H \bullet (A_{\wedge\tau})_t$ *for each $\tau \in \mathcal{T}$ and $t \in \mathbb{R}^+$.*

(iii) If a sequence $\{H_n\}$ in $\text{loc}\mathcal{H}_{\text{Bdd}}$ is locally uniformly bounded and converges pointwise (in t and ω) to 0 then $H_n \bullet A \xrightarrow{ucpc} 0$.

As you can see, there is really not much subtlety beyond the usual measure theory in the construction of stochastic integrals with respect to $\mathcal{FV}$-processes.

REMARK. The integral $H \bullet L_t$ can be defined even for processes that are not predictable or locally bounded. In fact, as there are no martingales involved in the construction, predictability is irrelevant. However, functions in $\text{loc}\mathcal{H}_{\text{Bdd}}$ will have stochastic integrals defined for both $\mathcal{FV}_0$-processes and $\text{loc}\mathcal{M}_0^2(\mathbb{R}^+)$-processes.

3. Stochastic integrals with respect to semimartingales

By combining the results from the previous Section with results from Project 6, we arrive at a most satisfactory definition of the stochastic integral for a very broad class of processes.

<3> **Definition.** *An R-process X is called a* ***semimartingale****, for a given standard filtration $\{\mathcal{F}_t : t \in \mathbb{R}^+\}$, if it can be decomposed as $X_t = X_0 + M_t + A_t$ with $M \in \text{loc}\mathcal{M}_0^2(\mathbb{R}^+)$ and $A \in \mathcal{FV}_0$. Write $\mathcal{SMG}$ for the class of all semimartingales and $\mathcal{SMG}_0$ for those semimartinagles with $X_0 \equiv 0$.*

$\mathcal{SMG}$ is nonstandard notation

Notice that $\mathcal{SMG}_0$ is stable under stopping. Moreover, every local semimartingale is a semimartingale, a fact that is surprisingly difficult (Dellacherie & Meyer 1982, §VII.26) to establish directly.

The stochastic integral $H \bullet X$ is defined as the sum of the stochastic integrals with respect to the components M and A. The value X_0 plays no role in this definition, so we may as well assume $X \in \mathcal{SMG}_0$. The resulting integral inherits the properties shared by integrals with respect to $\mathcal{FV}_0$ and integrals with respect to $\text{loc}\mathcal{M}_0^2(\mathbb{R}^+)$.

linear in thealmost sure sense

<4> **Theorem.** *For each X in $\mathcal{SMG}_0$, there is a linear map $H \mapsto H \bullet X$ from $\text{loc}\mathcal{H}_{\text{Bdd}}$ into $\mathcal{SMG}_0$ such that:*

(i) $((0,\tau]] \bullet X_t = X_{t\wedge\tau}$ *for each $\tau \in \mathcal{T}$ and $t \in \mathbb{R}^+$.*

(ii) $H \bullet X_{t\wedge\tau} = (H((0,\tau]]) \bullet X_t = H \bullet (X_{\wedge\tau})_t$ *for each $\tau \in \mathcal{T}$ and $t \in \mathbb{R}^+$.*

(iii) If a sequence $\{H_n\}$ in $\text{loc}\mathcal{H}_{\text{Bdd}}$ is locally uniformly bounded and converges pointwise (in t and ω) to 0 then $H_n \bullet X \xrightarrow{ucpc} 0$.

Conversely, let ψ be another linear map from $\text{loc}\mathcal{H}_{\text{Bdd}}$ into the set of R-processes having at least the weaker properties:

(iv) $\psi(((0,\tau]])_t = X_{t\wedge\tau}$ *almost surely, for each $\tau \in \mathcal{T}$ and $t \in \mathbb{R}^+$.*

(vi) If a sequence $\{H_n\}$ in $\text{loc}\mathcal{H}_{\text{Bdd}}$ is locally uniformly bounded and converges pointwise (in t and ω) to 0 then $\psi(H_n)_t \to 0$ in probability, for each fixed t.

Then $\psi(H)_t = H \bullet X_t$ almost surely for every t.

Remarks. The converse shows, in particular, that the stochastic integral $H \bullet X$ does not depend on the choice of the processes M and A in the semimartingale decomposition of X.

In general, I say that two processes X and Y are equal for ***almost all paths*** if $\mathbb{P}\{\exists t : X_t(\omega) \neq Y_t(\omega)\} = 0$. For processes with cadlag sample paths, this property is equivalent to $\mathbb{P}\{\omega : X_t(\omega) \neq Y_t(\omega)\} = 0$ for each t.

Outline of the proof of the converse. Define

$$\mathcal{H} := \{H \in \mathcal{H}_{\text{Bdd}} : \psi(H)_t = H \bullet X_t \text{ almost surely, for each } t \in \mathbb{R}^+\}$$

- Show that $((0,\tau]] \in \mathcal{H}$, for each $\tau \in \mathcal{T}$.
- Show that $\mathcal{H}$ is a λ-space. Hint: If $H_n \in \mathcal{H}$ and $H_n \uparrow H$, with H bounded, apply (iii) and (vi) to the uniformly bounded sequence $H - H_n$.
- Deduce that $\mathcal{H}$ equals $\mathcal{H}_{\text{Bdd}}$.
- Extend the conclusion to $\text{loc}\mathcal{H}_{\text{Bdd}}$. Hint: If $H \in \text{loc}\mathcal{H}_{\text{Bdd}}$, with $|H_{\wedge\tau_k}| \leq C_k$ for stopping times $\tau_k \uparrow \infty$, show that the processes $H_n := H((0,\tau_n]]$ are locally uniformly bounded and converge pointwise to H.

□

I have found the properties of the stochastic integral asserted by the Theorem to be adequate for many arguments. I consider it a mark of defeat if I have to argue separately for the $\text{loc}\mathcal{M}_0^2(\mathbb{R}^+)$ and $\mathcal{FV}_0$ cases to establish a general result about semimartingales. You might try Problem [3] or [4] for practice.

The class of semimartingales is quite large. It is stable under sums (not surprising) and products (very surprising—see the next Section) and under exotic things like change of measure (to be discussed in a later Project). Even more surprisingly, semimartingales are the natural class of integrators for stochastic integrals; they are the unexpected final product of a long sequence of ad hoc improvements. You might consult Protter (1990, pages 44; 87–88; 114), who expounded the whole theory by starting from plausible linearity and continuity assumptions then working towards the conclusion that only semimartingales can have the desired properties.

Characterization due to Dellacherie? Meyer? Bichteler? Métivier? Check history.

4. Quadratic variation

In the proof of Lévy's martingale characterization of Brownian Motion, you saw how a sum of squares of increments of Brownian motion, taken over a partition via stopping times of an interval $[0, t]$, converges in probability to t. In fact, if one allows random limits, the behaviour is a general property of semimartingales. The limit is called the ***quadratic variation process*** of the semimartingale.

It is easiest to establish existence of the quadratic variation by means of an indirect stochastic integral argument. Suppose X is an R-processes with $X_0 \equiv 0$. Define the left-limit process $X_t^{\ominus} := X(t-, \omega) := \lim_{s \uparrow\uparrow t} X(s, \omega)$. (Do we need to define $X_0^{\ominus}$?)

Awkward and nonstandard notation, $X^{\ominus}$, but I want X^- for the negative part of X.

- Show that $X^{\ominus} \in \text{loc}\mathcal{H}_{\text{Bdd}}$.

<5> **Definition.** *The quadratic variation process of an X in $\mathcal{SMG}_0$ is defined as $[X, X]_t := X_t^2 - 2(X^{\ominus} \bullet X)_t$ for $t \in \mathbb{R}^+$. For general $Z \in \mathcal{SMG}$, define $[Z, Z] := [X, X]$ where $X_t := Z_t - Z_0$.*

The logic behind the name *quadratic variation* and one of the main reasons for why it is a useful process both appear in the next Theorem. The first assertion of the Theorem could even be used to define quadratic variation, but then we would have to work harder to prove existence of the limit (as for the quadratic variation of Brownian motion).

<6> **Definition.** *A **random grid** $\mathbb{G}$ is defined by a finite sequence of finite stopping times $0 \le \tau_0 \le \tau_1 \le \ldots \le \tau_k$. The mesh of the grid is defined as* $\text{mesh}(\mathbb{G}) := \max_i |\tau_{i+1} - \tau_i|$*; the max of the grid is defined as* $\max(\mathbb{G}) := \tau_k$.

To avoid double subscripting, let me write $\sum_{\mathbb{G}}$ to mean a sum taken over the stopping times that make up $\mathbb{G}$.

<7> **Theorem.** *Suppose $X \in \mathcal{SMG}_0$ and $\{\mathbb{G}_n\}$ is a sequence of random grids with* $\text{mesh}(\mathbb{G}_n) \xrightarrow{a.s.} 0$ *and* $\max(\mathbb{G}_n) \xrightarrow{a.s.} \infty$. *Then:*

(i) $\sum_{\mathbb{G}_n} \left(X_{t\wedge\tau_{i+1}} - X_{t\wedge\tau_i}\right)^2 \xrightarrow{ucpc} [X, X]_t$.

(ii) The process $[X, X]$ has increasing sample paths;

(iii) If τ is a stopping time then $[X_{\wedge\tau}, X_{\wedge\tau}] = [X, X]_{\wedge\tau}$.

Mention jumps as well?

Outline of proof. Without loss of generality suppose $X_0 \equiv 0$. Consider first a fixed t and a fixed grid $\mathbb{G}$: $0 = \tau_0 \le \tau_1 \le \ldots \le \tau_k$.

- Define a left-continuous process $H_{\mathbb{G}} = \sum_{\mathbb{G}} X_{\tau_i}((\tau_i, \tau_{i+1}]]$. Show that $H \in \text{loc}\mathcal{H}_{\text{Bdd}}$ and
$$H_{\mathbb{G}} \bullet X_t = \sum_{\mathbb{G}} X_{\tau_i}\left(X_{t\wedge\tau_{i+1}} - X_{t\wedge\tau_i}\right)$$
Hint: Look at Problem [3].
- Except on a negligible set of paths (which I will ignore for the rest of the proof), show that $H_{\mathbb{G}}$ converges pointwise to the left-limit process $X^{\ominus}$ as $\text{mesh}(\mathbb{G}) \to 0$ and $\max(\mathbb{G}) \to \infty$. Show also that $\{H_{\mathbb{G}}\}$ is locally uniformly bounded. Hint: Consider stopping times $\sigma_k := \inf\{s : |X_s| \ge k\}$.

- Abuse notation by writing $\Delta_i X$ for $X_{t\wedge\tau_{i+1}} - X_{t\wedge\tau_i}$. Invoke the continuity property of the stochastic integral, along a sequence of grids with $\text{mesh}(\mathbb{G}_n) \to 0$ and $\max(\mathbb{G}_n) \to \infty$, to deduce that

$$\sum\nolimits_{\mathbb{G}_n} X_{\tau_i}(\Delta_i X) = H_{\mathbb{G}_n} \bullet X_t \xrightarrow{ucpc} X^{\ominus} \bullet X_t$$

- Show that

$$2H_{\mathbb{G}_n} \bullet X_t + \sum\nolimits_{\mathbb{G}_n} (\Delta_i X)^2 = X^2_{t\wedge\tau_k} \xrightarrow{ucpc} X_t^2.$$

- Complete the proof of (i).
- Establish (ii) by taking the limit along a sequence of grids (deterministic grids would suffice) for which both s and t are always grid points. Note: The sums of squared increments that converge to $[X, X]_t$ will always contain extra terms in addition to those for sums converging to $[X, X]_s$.
- For assertion (iii), merely note that $\tau \wedge t$ is one of the points in the interval $[0, t]$ over which the convergence in probability is uniform. Thus

$$\sum\nolimits_{\mathbb{G}_n} \left(X_{t\wedge\tau_{i+1}\wedge\tau} - X_{t\wedge\tau_i\wedge\tau}\right)^2 \xrightarrow{\mathbb{P}} [X, X]_{t\wedge\tau}.$$

Interpret the left-hand side as an approximating sum of squares for $[X_{\wedge\tau}, X_{\wedge\tau}]_t$. □

<8> **Corollary.** *The square of a semimartingale X is a semimartingale.*

Proof. Let $Z_t := X_t - X_0 = M_t + A_t$. Rearrange the definition of the square bracket process, $Z_t^2 = 2(Z^{\ominus} \bullet Z)_t + [Z, Z]_t$, to express Z_t^2 as a sum of a semimartingale and an increasing process. The process X_t^2 expands to $Z_t^2 + 2X_0 M_t + (2X_0 A_t + X_0^2)$.

- Show that the middle term is reduced to $\mathcal{M}_0^2(\mathbb{R}^+)$ by the stopping times $\tau_k \wedge \sigma_k$, where $\{\tau_k\}$ reduces M and $\sigma_k := 0\{|X_0| > k\} + \infty\{|X_0| \le k\}$. □

<9> **Corollary.** *The product of two semimartingales is a semimartingale.*

- Use the ***polarization identity***, $4XY = (X+Y)^2 - (X-Y)^2$, and the fact that sums of semimartingales are semimartingales, to reduce the assertion to the previous Corollary.

<10> **Definition.** *The **square bracket process** $[X, Y]$ of two semimartingales X and Y (also known as the quadratic covariation process of X and Y) is defined, by polarization, as*

$$4[X, Y] := [X+Y, X+Y] - [X-Y, X-Y].$$

If $X_0 \equiv 0$ and $Y_0 \equiv 0$ then $4[X, Y]_t$ equals

$$(X_t + Y_t)^2 - (X_t - Y_t)^2 - 2(X+Y)^{\ominus} \bullet (X+Y)_t + 2(X-Y)^{\ominus} \bullet (X-Y)_t$$

<11> $$= 4X_tY_t - 4X^{\ominus} \bullet Y_t - 4Y^{\ominus} \bullet X_t.$$

Remark. Notice that $[X, Y]$ is equal to the quadratic variation process $[X, X]$ when $X \equiv Y$. Notice also that $[X, Y] \in \mathcal{FV}_0$, being a difference of two increasing processes started at 0.

The square bracket process inherits many properties from the quadratic variation. For example, you might prove that a polarization argument derives the following result from Theorem <7>.

<12> **Theorem.** *Let X and Y be semimartingales, and $\{\mathbb{G}_n\}$ be a sequence of random grids with* $\text{mesh}(\mathbb{G}_n) \xrightarrow{a.s.} 0$ *and* $\max(\mathbb{G}_n) \xrightarrow{a.s.} \infty$. *Then*

<13> $$\sum\nolimits_{\mathbb{G}_n} \left(X_{t\wedge\tau_{i+1}} - X_{t\wedge\tau_i}\right)\left(Y_{t\wedge\tau_{i+1}} - Y_{t\wedge\tau_i}\right) \xrightarrow{ucp} [X,Y]_t,$$

and $[X_{\wedge\tau}, Y_{\wedge\tau}] = [X_{\wedge\tau}, Y] = [X, Y_{\wedge\tau}] = [X,Y]_{\wedge\tau}$ *for each stopping time* τ,

Problems

[1] Suppose $f = f_1 - f_2$, where f_1 and f_2 are increasing functions on $\mathbb{R}^+$. Show that

$$V_f[0,b] \le V_{f_1}[0,b] + V_{f_2}[0,b] = f_1(b) - f_1(0) + f_2(b) - f_2(0).$$

Deduce that f is of finite variation.

[2] Suppose f is a function on $\mathbb{R}^+$ with finite variation, in the sense of Section 1. Temorarily drop the subscript f on the variation functions.

(i) Suppose $\mathbb{G}$ is a grid on $[a,b]$ and that s is point of (a,b) that is not already a grid point. Show that $V(\mathbb{G},[a,b])$ is increased if we add s as a new grid point.

(ii) Show that $V[0,a] + V[a,b] = V[0,b]$ for all $a < b$. Deduce that $t \mapsto V[0,t]$ is an increasing function

(iii) Suppose $0 < s < t$. Show that

$$V[0,t] - f(t) = V[0,s] - f(s) + f(s) - f(t) + V[s,t] \ge V[0,s] - f(s).$$

Hint: Consider a two-point grid on $[s,t]$.

(iv) Now suppose f is right-continuous at some $a \in \mathbb{R}^+$. For a fixed $b > a$ and an $\epsilon > 0$ choose a grid

$$\mathbb{G}: \quad a = t_0 < t_1 < \ldots < t_N = b$$

for which $V(\mathbb{G},[a,b]) > V[a,b] - \epsilon$. With no loss of generality suppose $|f(t_1) - f(a)| < \epsilon$. Show that

$$\epsilon + V[t_1,b] \ge V(\mathbb{G},[a,b]) > V[a,t_1] + V[t_1,b] - \epsilon$$

Deduce that $t \mapsto V[0,t]$ is continuous from the right at a.

(v) If f is right-continuous, show that $V_f[a,b]$ can be determined by taking a supremum over equispaced grids on $[a,b]$.

(vi) If X is an R-processes with sample paths of finite variation, show that it can be expressed as the difference of two R-processes with increasing sample paths. [The issue is whether $V_{X(\cdot,\omega)}[0,t]$ is adapted.]

[3] Suppose σ and τ are stopping times and $X \in \mathcal{SMG}$. With Y an $\mathcal{F}_\sigma$-measurable random variable, define $H = Y(\omega)((\sigma,\tau]]$. Show that $H \bullet X_t = Y(\omega)\left(X_{t\wedge\tau} - X_{t\wedge\sigma}\right)$ by the following steps.

(i) Start with the case where $Y = F \in \mathcal{F}_\sigma$. Define new stopping times $\sigma' = \sigma F + \infty F^c$ and $\tau' := \tau F + \infty F^c$. Show that

$$(F((\sigma,\tau]]) \bullet X_t = X_{t\wedge\tau'} - X_{t\wedge\sigma'} = F\left(X_{t\wedge\tau} - X_{t\wedge\sigma}\right).$$

(ii) Extend the equality to all bounded, $\mathcal{F}_\sigma$-measurable Y by a generating class argument.

(iii) For unbounded Y, define $H_n := Y\{|Y| \le n\}((\sigma,\tau]]$. Show that the sequence $\{H_n\}$ is locally uniformly bounded and it converges pointwise to H.

(iv) Complete the argument.

[4] If H and K are in loc$\mathcal{H}_{\text{Bdd}}$, and X is a semimartingale, show that $K \bullet (H \bullet X) = (KH) \bullet X$ for almost all paths. Hint: For fixed H, define $\psi(K) := (HK) \bullet M$. What do you get when $K = ((0, \tau]]$?

[5] Suppose $H, K \in \text{loc}\mathcal{H}_{\text{Bdd}}$ and $X, Y \in \mathcal{SMG}_0$. Show that $[H \bullet X, K \bullet Y] = (HK) \bullet [X, Y]$ by the following steps.

(i) Consider first the case where $K \equiv 1$. Show that $H \mapsto [H \bullet X, Y]$ and $H \mapsto H \bullet [X, Y]$ are both linear maps from loc$\mathcal{H}_{\text{Bdd}}$ into $\mathcal{SMG}$, which agree when $H = ((0, \tau]]$.

(ii) Use a λ-space argument followed by a localization to extend the result to loc$\mathcal{H}_{\text{Bdd}}$.

(iii) Invoke part (ii)—or trivial rearrangements thereof—twice to transform to an iterated stochastic integral.

$$[H \bullet X, K \bullet Y] = H \bullet [X, K \bullet Y] = H \bullet \big(K \bullet [X, Y]\big).$$

□ (iv) Invoke Problem [4] to complete the argument.

[6] Suppose $M \in \text{loc}\mathcal{M}_0^2(\mathbb{R}^+)$.

(i) Show that the process $X_t := M_t^2 - [M, M]_t$ belongs to loc$\mathcal{M}_0^2(\mathbb{R}^+)$.

(ii) Suppose M has continuous sample paths and $[M, M]_t \equiv t$. Show that M is a standard Brownian motion.

References

Dellacherie, C. & Meyer, P. A. (1978), *Probabilities and Potential*, North-Holland, Amsterdam. (First of three volumes).

Dellacherie, C. & Meyer, P. A. (1982), *Probabilities and Potential B: Theory of Martingales*, North-Holland, Amsterdam.

Protter, P. (1990), *Stochastic Integration and Differential Equations*, Springer, New York.

PROJECT 8

$\mathcal{SMG}$ is nonstandard notation

An R-process X is called a ***semimartingale***, for a given standard filtration $\{\mathcal{F}_t : t \in \mathbb{R}^+\}$, if it can be decomposed as $X_t = X_0 + M_t + A_t$ with $M \in \text{loc}\mathcal{M}_0^2(\mathbb{R}^+)$ and $A \in \mathcal{FV}_0$. Write $\mathcal{SMG}$ for the class of all semimartingales and $\mathcal{SMG}_0$ for those semimartinagles with $X_0 \equiv 0$.

1. Corrections

In my enthusiasm for a single definition of localization, which could be applied to both $\mathcal{M}_0^2(\mathbb{R}^+)$ and $\mathcal{H}_{\text{Bdd}}$, I created an awkward problem for processes H defined only on $(0, \infty) \times \Omega$. If $H_0(\omega)$ is not defined, what does $H(t \wedge \tau(\omega), \omega)$ mean at those ω for which $\tau(\omega) = 0$? It would be much better to follow traditional and defineloc$\mathcal{H}_{\text{Bdd}}$ to consist of those predictable processes H for which there exist stopping times $\tau_k \uparrow \infty$ and finite constants C_k such that

$$|H((0, \tau_k]]| \le C_k \qquad \text{for each } k.$$

Notice that there are no longer problems at ω for which $\tau_k(\omega) = 0$, because $\{t \in \mathbb{R}^+ : 0 < t \le \tau_k(\omega)\} = \emptyset$ for such ω.

Similarly, I should have defined local uniform boundedness of a sequence $\{H_n\}$ in loc$\mathcal{H}_{\text{Bdd}}$ to mean existence of stopping times $\tau_k \uparrow \infty$ and finite constants C_k such that

$$|H_n((0, \tau_k]]| \le C_k \qquad \text{for each } n \text{ and } k.$$

I was also too vague about the definition of L-processes on $\mathbb{R}^+ \times \Omega$. Should such a process X be defined at $t = 0$? Should we require existence of a finite right-hand limit at $t = 0$? Should we require existence of a limit as $t \to \infty$? To make sense of my assertion that L-processes belong to loc$\mathcal{H}_{\text{Bdd}}$, I should regard X as an adapted process defined on $(0, \infty) \times \Omega$ with sample paths that are left-continuous on $(0, \infty)$, with no assumptions about the behavior as $t \to \infty$. I also need existence of a finite right limit at each t in $[0, \infty)$. With these assumptions, the stopping times

$$\tau_k := \inf\{t \in \mathbb{R}^+ : |X_t| > k\}$$

have the property that $|X((0, \tau_k]]| \le k$. Also we have $\tau_k \uparrow \infty$, because $X(\cdot, \omega)$ is bounded on each bounded interval $(0, M]$: You need a compactness argument to get a covering of $[0, M]$ by finitely many intervals $(t_i - \delta_i, t_i + \delta_i)$ within which

$$\begin{aligned} |X(t, \omega) - X(t_i, \omega)| &\le \epsilon && \text{for } t_i - \delta_i < t \le t_i \\ |X(t, \omega) - X(t_i+, \omega)| &\le \epsilon && \text{for } t_i < t \le t_i + \delta_i. \end{aligned}$$

I have also decided that it would be better to slightly change parts (iii) and (vi) of the basic theorem about semimartinagles, to simplify one step in the typical generating class argument.

linear in almost sure sense

<1> **Theorem.** *For each X in $\mathcal{SMG}_0$, there is a linear map $H \mapsto H \bullet X$ from* $\text{loc}\mathcal{H}_{\text{Bdd}}$ *into $\mathcal{SMG}_0$ such that:*

(i) $((0,\tau]] \bullet X_t = X_{t\wedge\tau}$ *for each $\tau \in \mathcal{T}$ and $t \in \mathbb{R}^+$.*

(ii) $H \bullet X_{t\wedge\tau} = (H((0,\tau]]) \bullet X_t = H \bullet (X_{\wedge\tau})_t$ *for each $\tau \in \mathcal{T}$ and $t \in \mathbb{R}^+$.*

(iii) If a sequence $\{H^{(n)} : n \in \mathbb{N}\} \subseteq \text{loc}\mathcal{H}_{\text{Bdd}}$ is locally uniformly bounded and $H^{(n)}(t,\omega) \to H(t,\omega)$ for each (t,ω), then $H \in \text{loc}\mathcal{H}_{\text{Bdd}}$ and $H^{(n)} \bullet X \xrightarrow{ucpc} H \bullet X$.

Conversely, let ψ be another linear map from $\text{loc}\mathcal{H}_{\text{Bdd}}$ into $\mathcal{SMG}_0$ having at least the weaker properties:

(iv) $\psi(((0,\tau]])_t = X_{t\wedge\tau}$ almost surely, for each $\tau \in \mathcal{T}$ and $t \in \mathbb{R}^+$.

(vi) If a sequence $\{H^{(n)} : n \in \mathbb{N}\} \subseteq \text{loc}\mathcal{H}_{\text{Bdd}}$ is locally uniformly bounded and $H^{(n)}(t,\omega) \to H(t,\omega)$ for each (t,ω), then $\psi(H^{(n)})_t \to \psi(H)_t$ in probability, for each fixed t.

Then $\psi(H)_t = H \bullet X_t$ almost surely for every t.

REMARK. I did attempt to weaken the pointwise convergence assumptions in (iii) and (vi) to: $H_t^{(n)} \to H_t$ almost surely for each t. Unfortunately, this change complicates (invalidates?) the argument that $H \in \text{loc}\mathcal{H}_{\text{Bdd}}$. I do not know whether it is worthwhile attempting such a modification.

2. Quadratic variation

From Project 7:

<2> **Definition.** *The quadratic variation process of an X in $\mathcal{SMG}_0$ is defined as $[X,X]_t := X_t^2 - 2(X^{\ominus} \bullet X)_t$ for $t \in \mathbb{R}^+$. For general $Z \in \mathcal{SMG}$, define $Z, Z] := [X,X]$ where $X_t := Z_t - Z_0$.*

<3> **Definition.** *A **random grid** $\mathbb{G}$ is defined by a finite sequence of finite stopping times $0 \le \tau_0 \le \tau_1 \le \ldots \le \tau_k$. The mesh of the grid is defined as* $\text{mesh}(\mathbb{G}) := \max_i |\tau_{i+1} - \tau_i|$*; the max of the grid is defined as* $\max(\mathbb{G}) := \tau_k$.

<4> **Theorem.** *Suppose $X \in \mathcal{SMG}$ and $\{\mathbb{G}_n\}$ is a sequence of random grids with* $\text{mesh}(\mathbb{G}_n) \xrightarrow{a.s.} 0$ *and* $\max(\mathbb{G}_n) \xrightarrow{a.s.} \infty$. *Then:*

(i) $\sum_{\mathbb{G}_n} \left(X_{t\wedge\tau_{i+1}} - X_{t\wedge\tau_i}\right)^2 \xrightarrow{ucpc} [X,X]_t$.

(ii) The process $[X,X]$ has increasing sample paths;

Mention jumps as well?

(iii) If τ is a stopping time then $[X_{\wedge\tau}, X_{\wedge\tau}] = [X,X]_{\wedge\tau}$.

REMARK. It would perhaps be cleaner to assume $\text{mesh}(\mathbb{G}_n) \to 0$ and $\max(\mathbb{G}_n) \to \infty$ for every ω, to fit with the pointwise convergence assumptions in Theorem <1>. This effect could also be achieved by changing each τ_k on a negligible set. For a standard filtration, the change could be made without disturbing any measurability assumptions.

The ***square bracket process*** $[X,Y]$ of two semimartingales X and Y (also known as the quadratic covariation process of X and Y) is defined, by polarization, as

$$4[X,Y] := [X+Y, X+Y] - [X-Y, X-Y].$$

If $X_0 \equiv 0$ and $Y_0 \equiv 0$ then $4[X,Y]_t$ equals

$$(X_t+Y_t)^2 - (X_t-Y_t)^2 - 2(X+Y)^{\ominus} \bullet (X+Y)_t + 2(X-Y)^{\ominus} \bullet (X-Y)_t$$

<5> $$= 4X_tY_t - 4X^{\ominus} \bullet Y_t - 4Y^{\ominus} \bullet X_t.$$

Notice that $[X, Y]$ is equal to the quadratic variation process $[X, X]$ when $X \equiv Y$. Notice also that $[X, Y] \in \mathcal{FV}_0$, being a difference of two increasing processes started at 0.

The square bracket process inherits many properties from the quadratic variation. For example, a polarization argument derives the following result from Theorem <4>.

<6> **Theorem.** *Suppose* $X, Y \in \mathcal{SMG}$ *and* $\{\mathbb{G}_n\}$ *is a sequence of random grids with* $\mathrm{mesh}(\mathbb{G}_n) \xrightarrow{a.s.} 0$ *and* $\max(\mathbb{G}_n) \xrightarrow{a.s.} \infty$. *Then*

<7>
$$\sum\nolimits_{\mathbb{G}_n} \left(X_{t\wedge\tau_{i+1}} - X_{t\wedge\tau_i}\right)\left(Y_{t\wedge\tau_{i+1}} - Y_{t\wedge\tau_i}\right) \xrightarrow{ucpc} [X, Y]_t,$$

and $[X_{\wedge\tau}, Y_{\wedge\tau}] = [X_{\wedge\tau}, Y] = [X, Y_{\wedge\tau}] = [X, Y]_{\wedge\tau}$ *for each stopping time* τ,

I am slightly suspicious of the following argument.

Outline of last part of the proof. Temporarily write $W_n(t, X, Y)$ for the sum on the left-hand side of <7>.

- Show that there is no loss of generality in assuming that τ is one of the grid points. Hint: Consider a new grid
$$0 = \tau_0 \wedge \tau \le \tau_1 \wedge \tau \le \ldots \le \tau_k \wedge \tau \le \tau \le \tau \vee \tau_1 \le \ldots \le \tau \vee \tau_k$$
- For the grid $\mathbb{G}n$, suppose $\tau = \tau_\ell$. Show that
$$\begin{aligned} W_n(t\wedge\tau, X, Y) &= \sum\nolimits_{i=0}^{\ell-1} \left(X_{t\wedge\tau_{i+1}} - X_{t\wedge\tau_i}\right)\left(Y_{t\wedge\tau_{i+1}} - Y_{t\wedge\tau_i}\right) \\ &= W_n(t, X_{\wedge\tau}, Y) = W_n(t, X, Y_{\wedge\tau}) = W_n(t, X_{\wedge\tau}, Y_{\wedge\tau}) \end{aligned}$$
- Invoke uniform convergence over $[0, t]$ in probability. □

3. Itô formulae

Treat processes with jumps, or just cite Dellacherie & Meyer (1982, §VIII.24–28) or Protter (1990, page 71)?

Suppose X and Y are semimartingales with continuous paths, such that the two-dimensional random process $\{(X_t, Y_t) : t \in \mathbb{R}^+\}$ takes values in an open subset G of $\mathbb{R}^2$. Suppose Y has paths of bounded variation.

Let f be a continuous, real-valued function on G with two continuous partial derivatives f_x and f_{xx} with respect to its first argument and a continuous partial derivative f_y with respect to its second argument.

Define new processes by

$$\begin{aligned} F_x(s,\omega) &:= f_x\big(X(s,\omega), Y(s,\omega)\big), \\ F_{xx}(s,\omega) &:= f_{xx}\big(X(s,\omega), Y(s,\omega)\big) \\ F_y(s,\omega) &:= f_y\big(X(s,\omega), Y(s,\omega)\big). \end{aligned}$$

Each of them is adapted and has continuous paths; each process is predictable.

<8> **Itô Formula.** *The process* $f(X_s, Y_s)$ *is a semimartingale with*
$$f(X_t, Y_t) - f(X_0, Y_0) = (F_x \bullet X)_t + {}^1\!/_2\left(F_{xx} \bullet [X, X]\right)_t + (F_y \bullet Y)_t$$
for each t *in* $\mathbb{R}^+$.

REMARK. The Itô formula is often written in the suggestive form
$$df(X_t, Y_t) = f_x(X_t, Y_t)\,dX_t + {}^1\!/_2 f_{xx}(X_t, Y_t)\,d[X, X]_t + f_y(X_t, Y_t)\,dY_t,$$
which hints at its origins as a sum of small increments.

Proof. Let K be a compact subset of G. Define
$$\sigma := \inf\{t \in \mathbb{R}^+ : (X_t, Y_t) \notin K\}.$$

I was tired when sketching the proof. Beware of stupidities.

Replace X and Y by the corresponding stopped processes $X_{\wedge\sigma}$ and $Y_{\wedge\sigma}$.

- Show that the formula is trivially true for the stopped processes if $(X_0, Y_0) \notin K$.
- For each $\epsilon > 0$ show that there exists a $\delta > 0$ for which: if $(x, y) \in K$ and $\max(|\Delta x|, |\Delta y|) \le \delta$ then

$$|f_{\Box}(x + \Delta x, y + \Delta y) - f_{\Box}(x, y)| \le \epsilon \qquad \text{where } \Box = x \text{ or } xx \text{ or } y.$$

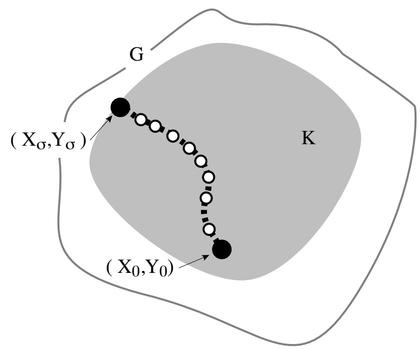
- For $\max(|\Delta x|, |\Delta y|) \le \delta$ and $(x, y) \in K$, show that

$$\begin{aligned} &f(x+\Delta x, y + \Delta y) - f(x, y) \\ &\quad = (\Delta x) f_x(x, y) + 1/2(\Delta x)^2 f_{xx}(x, y) + (\Delta y) f_y(x, y) + \text{REM} \\ &\qquad \text{where REM} \le \epsilon\big(1/2(\Delta x)^2 + |\Delta y|\big) \end{aligned}$$

- Fix t. Let δ_n correspond to some sequence $\epsilon_n \downarrow 0$. Define a grid $\mathbb{G}_n$ via stopping times

$$\tau_{i+1} := \inf\{s \ge \tau_i : |(X, Y)_s - (X, Y)_{\tau_i}| \ge \delta_n\} \wedge t \wedge \sigma.$$

Show that there exist integers $k(n)$ such that $\mathbb{P}\{\tau_{k(n)} = t \wedge \sigma\} \to 1$ as $n \to \infty$.



- Write $\Delta_i X$ for $X_{\tau_{i+1}} - X_{\tau_i}$, and similarly for Y. Show that $f(X_{\tau_{k(n)}}, Y_{\tau_{k(n)}}) - f(X_0, Y_0)$ differs from

<9> $$\sum_{i=0}^{k(n)-1} (\Delta_i X) F_x(\tau_i) + 1/2(\Delta_i X)^2 F_{xx}(\tau_i) + (\Delta_i Y) F_y(\tau_i)$$

by a quantity that tends in probability to zero.

- Show that the contribution from the first summand in <9> equals $(H_n \bullet X)_t$, where

$$H_n(s, \omega) = \sum_{i=0}^{k(n)-1} F_x(\tau_i, \omega)((\tau_i . \tau_{i+1}]],$$

which is uniformly bounded and converges pointwise to F_x.

- Deduce that

$$\sum_{i=0}^{k(n)-1} (\Delta_i X) F_x(\tau_i) \xrightarrow{ucpc} F_x \bullet X_{t\wedge\sigma}$$

- Argue similarly for the contribution from the third summand in <9> .
- Define $Z_t := X_t - X_0$. Abbreviate $Z_{\tau_{i+1}} - Z_{\tau_i}$ to $\Delta_i Z$. Show that

cf. Protter (1990, page 69)

$$\begin{aligned} \sum_{i=0}^{k(n)-1} (\Delta_i X)^2 F_{xx}(\tau_i) &= \sum_{i=0}^{k(n)-1} F_{xx}(\tau_i)(Z^2_{\tau_{i+1}} - Z^2_{\tau_i}) \\ &\quad - 2\sum_{i=0}^{k(n)-1} \big(F_{xx}(\tau_i) Z_{\tau_i}\big)(\Delta_i Z). \end{aligned}$$

- Show that the right-hand side converges in probability to

$$\begin{aligned} F_{xx} \bullet Z^2_{t\wedge\sigma} - 2(F_{xx} Z) \bullet Z_{t\wedge\sigma} &= F_{xx} \bullet (Z^2 - 2Z \bullet Z)_{t\wedge\sigma} \\ &= F_{xx} \bullet [Z, Z]_{t\wedge\sigma} = F_{xx} \bullet [X, X]_{t\wedge\sigma}. \end{aligned}$$

- Deduce that

$$
\begin{aligned}
&f(X_{t\wedge\sigma}, Y_{t\wedge\sigma}) - f(X_0, Y_0)\\
&\qquad = (F_x \bullet X_{\wedge\sigma})_t + {}^1\!/\!_2(F_{xx} \bullet [X_{\wedge\sigma}, X_{\wedge\sigma}])_t + (F_y \bullet Y_{\wedge\sigma})_t\\
&\qquad = (F_x \bullet X)_{t\wedge\sigma} + {}^1\!/\!_2(F_{xx} \bullet [X, X])_{t\wedge\sigma} + (F_y \bullet Y)_{t\wedge\sigma}.
\end{aligned}
$$

□ • Complete the proof by letting K expand up to G, so that $\sigma \uparrow \infty$.

REMARKS.

(i) There would be nothing to gain by requiring existence of second-order partial derivatives f_{xy} and f_{yy}, because the corresponding bracket process $[X, Y]$ and $[Y, Y]$ are both zero—the process Y has paths of finite variation.

(ii) The process ${}^1\!/\!_2 F_{xx} \bullet [X, X] + F_y \bullet Y$ is in $\mathcal{FV}$. If $X \in \text{loc}\mathcal{M}_0^2(\mathbb{R}^+)$ then $F_x \bullet X \in \text{loc}\mathcal{M}_0^2(\mathbb{R}^+)$. The Itô formula then gives the semimartingale decomposition for the process $f(X_t, Y_t)$.

The story in Remark (i) changes if Y does not have paths of bounded variation. The error term $\epsilon_n \sum_i |\Delta_i Y|$ would no longer disappear in the limit. We would instead need continuous second order partial derivatives f_{xy} and f_{yy} to handle the contributions from the $\Delta_i Y$ increments to the Taylor expansion (to quadratic terms) in both variables. Error terms like

$$\epsilon_n \sum\nolimits_i (\Delta_i Y)^2 + (\Delta_i X)(\Delta_i Y)$$

would again converge in probability to zero. The cross-product term

$$
\begin{aligned}
&\sum\nolimits_i F_{xy}(\tau_i)(\Delta_i X)(\Delta_i Y)\\
&\quad = \sum\nolimits_i F_{xy}(\tau_i)(X_{\tau_{i+1}} Y_{\tau_{i+1}} - X_{\tau_i} Y_{\tau_i}) - \sum\nolimits_i F_{xy}(\tau_i)\left(X_{\tau_i}(\Delta_i Y) + Y_{\tau_i}(\Delta_i X)\right)
\end{aligned}
$$

would converge in probability to

$$F_{xy} \bullet (XY - X_0 Y_0 - X \bullet Y - Y \bullet X)_t = F_{xy} \bullet [X, Y]_t.$$

A similar argument works for functions of more than two semimartingales.

<10> **Multiprocess Itô Formula.** *Suppose $X^{(1)}, \dots X^{(d)}$ and $Y^{(1)}, \dots Y^{(d')}$ are semimartingales with continuous paths, such that the $d + d'$-dimensional random process $(\mathbf{X}, \mathbf{Y})$ takes values in an open subset G of $\mathbb{R}^{d+d'}$. Suppose each $Y^{(\gamma)}$ has paths of finite variation.*

If f is a continuous, real-valued function on G with continuous partial derivatives $f_{x(\alpha)}$, $f_{x(\alpha),x(\beta)}$, $f_{y(\gamma)}$ for $\alpha, \beta = 1, \dots, d$ and $\gamma = 1, \dots, d'$, then $f(\mathbf{X}, \mathbf{Y})$ is a semimartingale with

$$
\begin{aligned}
f(\mathbf{X}_t, \mathbf{Y}_t) - f(\mathbf{X}_0, \mathbf{Y}_0) &= \sum\nolimits_\alpha F_{x(\alpha)} \bullet X_t^{(\alpha)}\\
&\quad + \sum\nolimits_\gamma F_{y(\gamma)} \bullet Y_t^{(\gamma)} + {}^1\!/\!_2 \sum\nolimits_{\alpha,\beta} F_{x(\alpha),x(\beta)} \bullet [X^{(\alpha)}, X^{(\beta)}]_t
\end{aligned}
$$

for each t in $\mathbb{R}^+$.

<11> **Example.** Let $\{X_t : t \in \mathbb{R}^+\}$ be a locally square integrable martingale with continuous sample paths. Its quadratic variation process $Y := [X, X]$ is continuous (invoke the ucpc of the sum of squared increments) and of bounded variation. To be on the safe side, let me also assume that $X_0 \equiv 0$, even though it is not necessary.

Remove assumption on X_0.

The semimartingale $Z_t := \exp(X_t - {}^1\!/\!_2 Y_t)$ is a candidate for an application of the Itô formula, with $f(x, y) = \exp(x - {}^1\!/\!_2 y)$. We have $F_x = F_{xx} = -2F_y = Z$, and

$$Z_t - Z_0 = Z \bullet X + {}^1\!/\!_2 Z \bullet [X, X]_t - {}^1\!/\!_2 Z \bullet Y_t = Z \bullet X_t.$$

The Z process is also a locally square integrable martingale with continuous
□ paths.

4. Problems

[1] Show that

$$[X_1 + X_2, Y_1 + Y_2] = [X_1, Y_1] + [X_1, Y_2] + [X_2, Y_1] + [X_2, Y_2],$$

for semimartingales X_1, X_2, Y_1, and Y_2.

[2] Suppose $X \in \mathcal{SMG}$ and $Y \in \mathcal{FV}$. Suppose that X has continuous sample paths. Show that $[X, Y]_t = 0$ almost surely, for each t. Hint: Consider a random grid defined by

$$\tau_{i+1} := (\tau_i + n^{-1}) \wedge \min\{t \geq \tau_i : |X(t) - X(\tau_i)| \geq n^{-1}\}.$$

[3] For H_1, H_2, K_1, K_2 in $\text{loc}\mathcal{H}^\infty$, and X_1, X_2, Y_1, Y_2 in $\mathcal{SMG}$, show that

$$\begin{aligned}
&[H_1 \bullet X_1 + K_1 \bullet Y_1, H_2 \bullet X_2 + K_2 \bullet Y_2] \\
&\quad = (H_1 H_2) \bullet [X_1, X_2] + (H_1 K_2) \bullet [X_1, Y_2] \\
&\quad + (K_1 H_2) \bullet [Y_1, X_2] + (K_1 K_2) \bullet [Y_1, Y_2].
\end{aligned}$$

[4] If $X \in \mathcal{SMG}$ has continuous paths, show that $[X, X]$ also has conntinuous paths.

References

Dellacherie, C. & Meyer, P. A. (1982), *Probabilities and Potential B: Theory of Martingales*, North-Holland, Amsterdam.

Protter, P. (1990), *Stochastic Integration and Differential Equations*, Springer, New York.

Project 9

> **Itô Formula:** Two-dimensional semimartingale (X, Y) with continuous paths and $Y \in \mathcal{FV}$, which ensures $[Y, Y] = 0$. Continuous, real-valued function f with enough continuous partial derivatives to define predictable processes
> $$F_x(s,\omega) := f_x\big(X(s,\omega), Y(s,\omega)\big),$$
> $$F_{xx}(s,\omega) := f_{xx}\big(X(s,\omega), Y(s,\omega)\big)$$
> $$F_y(s,\omega) := f_y\big(X(s,\omega), Y(s,\omega)\big).$$
> with continuous sample paths. Then $f(X_s, Y_s)$ is a semimartingale with
> $$f(X_t, Y_t) - f(X_0, Y_0) = (F_x \bullet X)_t + 1/2\big(F_{xx} \bullet [X, X]\big)_t + (F_y \bullet Y)_t$$
> for each t in $\mathbb{R}^+$.

1. Corrections

On Project 7 I asked you to show that

<1> $$[H \bullet X, Y] = H \bullet [X, Y] \qquad \text{for } X, Y \in \mathcal{SMG}_0 \text{ and } H \in \text{loc}\mathcal{H}_{\text{Bdd}}.$$

I implied that the proof was just a simple example of a generating class argument. As some of you discovered, the proof is a little more delicate. A clean argument can be extracted from ideas used by Protter (1990, section II.6).

<2> **Lemma.** *Suppose $\{H_n : n \in \mathbb{N}\} \subseteq \text{loc}\mathcal{H}_{\text{Bdd}}$ is locally uniformly bounded and $H_n \xrightarrow{ucpc} 0$. Suppose also that $Y \in \mathcal{SMG}$. Then $H_n \bullet Y \xrightarrow{ucpc} 0$.*

Proof. Suppose there is a t for which $\sup_{s\le t} |H_n \bullet Y_s|$ does not converge to zero in probability. For some $\epsilon > 0$ there is a subsequence along which $\mathbb{P}\{\sup_{s\le t} |H_n \bullet Y_s| > \epsilon\} > \epsilon$. Along a subsubsequence we have the same inequality as well as $\sup_{s\le t} |H_n(s)| \to 0$ almost surely; along the subsubsequence $\sup_{s\le t} |\{\omega \in N^c\}H_n(s,\omega)| \to 0$ for every ω, for some negligible set N. The sequence $K_n := \{\omega \in N^c\}H_n((0,t]]$ is locally uniformly bounded and it converges pointwise to zero. Each K_n is $\mathcal{P}$-measurable, because $(0,1] \times N \in \mathcal{P}$, and

$$\mathbb{P}\{\sup_{s\le t} |K_n \bullet Y_s - H_n \bullet Y_s| \ne 0\} = 0.$$

Along the subsubsequence, $K_n \bullet Y \xrightarrow{ucpc} 0$, which contradicts the property defining the first subsequence. □

To establish assertion <1>, consider the linear map

$$\begin{aligned}\psi(H) &:= [H \bullet X, Y] - H \bullet [X, Y]\\ &= (H \bullet X)Y - (H \bullet X)^{\ominus} \bullet Y - Y^{\ominus} \bullet (H \bullet X)\\ &\qquad - H \bullet \big(XY - X^{\ominus} \bullet Y - Y^{\ominus} \bullet X\big)\\ &= (H \bullet X)Y - (H \bullet X)^{\ominus} \bullet Y - H \bullet \big(XY - X^{\ominus} \bullet Y\big)\end{aligned}$$

because $Y^{\ominus} \bullet (H \bullet X) = (Y^{\ominus}H) \bullet X = H \bullet \big(Y^{\ominus} \bullet X\big)$.

You can check that $\psi((0,\tau]] = 0$ for $\tau \in \mathcal{T}$ and that $\psi(H_n - H) \xrightarrow{ucpc} 0$ if $\{H_n : n \in \mathbb{N}\}$ is locally uniformly bounded and $H_n \to H$ pointwise. I think the rest of the argument is routine.

Please inform me if you find more gaps in the proof.

2. Exponential martingales

Suppose $M \in \text{loc}\mathcal{M}_0^2(\mathbb{R}^+)$ has continuous sample paths. For $H \in \text{loc}\mathcal{H}_{\text{Bdd}}$, define

<3>
$$Z_t = \exp\left(iH \bullet M_t + \tfrac{1}{2}H^2 \bullet [M, M]_t\right)$$

- Invoke the complex analog of the Itô formula (or apply the result to real and imaginary parts) to show that

$$\begin{aligned} Z_t - 1 &= iZ \bullet (H \bullet M)_t - \tfrac{1}{2}Z \bullet [H \bullet M, H \bullet M]_t + \tfrac{1}{2}Z \bullet \left(H^2 \bullet [M, M]\right)_t \\ &= i(ZH) \bullet M_t \end{aligned}$$

  You may use any of the properties established in the problems for Project 8.
- Deduce that $Z - 1 \in \text{loc}\mathcal{M}_0^2(\mathbb{R}^+)$.

3. Lévy again

Define $\mathcal{U}_t = t$. Suppose $M \in \text{loc}\mathcal{M}_0^2(\mathbb{R}^+)$ with continuous sample paths and $M^2 - \mathcal{U} \in \text{loc}\mathcal{M}_0^2(\mathbb{R}^+)$. Show that M is a standard Brownian motion.

- Use Problem [2] to explain why $[M, M] = \mathcal{U}$.
- For fixed constants $0 = t_0 < t_1 < \ldots < t_\ell < \infty$ and real numbers $\{\theta_j\}$ define

$$H = \sum\nolimits_{j=0}^{\ell-1} \theta_j((t_j, t_{j+1}]]$$

  Show that, for $t \geq t_\ell$,

$$\begin{aligned} H \bullet M_t &= \sum\nolimits_j \theta_j \Delta_j M \qquad \text{where } \Delta_j M := M(t_{j+1}) - M(t_j) \\ H^2 \bullet \mathcal{U}_t &= \sum\nolimits_j \theta_j^2 \delta_j \qquad \text{where } \delta_j := t_{j+1} - t_j \end{aligned}$$

- For Z as in <3> and H as above, show that there is a sequence of stopping times $\tau_k \uparrow \infty$ for which $\mathbb{P}Z_{t \wedge \tau_k} = 1$ for all k.
- Invoke Dominated Convergence to deduce that

$$\mathbb{P}\exp\left(i \sum\nolimits_j \theta_j \Delta_j M\right) = \exp\left(-\tfrac{1}{2}\sum\nolimits_j \theta_j^2 \delta_j)\right)$$

- Conclude that M is a standard Brownian motion.

4. Brownian filtrations

Let $\{B(t, \omega : 0 \leq t \leq 1\}$ be a Brownian motion with continuous sample paths on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The ***Brownian filtration*** on Ω is defined by $\mathcal{F}_t := \sigma\left(\{B_s : 0 \leq s \leq t\} \cup \mathcal{N}\right)$, where $\mathcal{N}$ denotes the class of all $\mathbb{P}$-negligible sets. The submartingale B^2 has Doléans measure $\mu = \mathfrak{m} \otimes \mathbb{P}$.

<4> **Definition.** *A cadlag process $\{M_t : 0 \leq t \leq 1\}$ is said to be a local martingale if there exist stopping times $\tau_k \uparrow \infty$ for which each $M_{\wedge \tau_k}$ is a martingale.*

Maybe better to restrict definition to cases where $M_0 = 0$.

Local martingales (with $M_0 = 0$) with respect to the Brownian filtration have two striking properties:

(i) They have continuous sample paths. Thus they all belong to $\text{loc}\mathcal{M}_0^2[0, 1]$.

(ii) They can be represented as stochastic integrals.

See Problem [4] for the first assertion. The second will follow via an argument based on the Itô formula.

<5> **Theorem.** *For each* $X \in \mathcal{L}^2(\Omega, \mathcal{F}_1, \mathbb{P})$ *there exists an* $H \in \mathcal{L}^2(\mu)$ *such that* $X - \mathbb{P}X = H \bullet B_1$ *almost surely.*

Sketch of proof. Without loss of generality, suppose $\mathbb{P}X = 0$.

- Show that $\mathcal{R} := \{H \bullet B_1 : H \in L^2(\mu)\}$ is a closed vector subspace of $\mathcal{L}^2(\mathbb{P}, \mathcal{F}_1)$. Hint: If $H_n \bullet B_1 \to Y$ in $\mathcal{L}^2(\mathbb{P})$-norm, show that $\{H_n\}$ is a Cauchy sequence with a limit H in $\mathcal{L}^2(\mu_1)$. Deduce that $Y = H \bullet B_1$.
- Let Z denote the component of X that is orthogonal to $\mathcal{R}$. That is, $X = Z + K \bullet B_1$ for some $K \in \mathcal{L}^2(\mu)$ and $\mathbb{P}Z(H \bullet B)_1 = 0$ for all H in $\mathcal{L}^2(\mu)$. Show that $\mathbb{P}Z = 0$.
- Explain why we need to prove $Z = 0$ almost surely.
- Explain why it suffices to show $\mathbb{P}Zf(B) = 0$ for all bounded, $\mathcal{C}$-measurable functionals f on $C[0, 1]$.
- Explain why it suffices to consider functionals f that depend on B only through its values at a finite set of times.
- Explain why it suffices to consider functionals f that depend on B only through its increments $Y_j = B_{t_{j+1}} - B_{t_j}$ for a fixed set of times $0 = t_0 < t_1 < \ldots < t_k = 1$. That is, why is it enough to prove $\mathbb{P}Zg(\mathbf{Y}) = 0$ for all bounded, measurable functions g on $\mathbb{R}^k$?
- Invoke Problem [3] to show that it is enough to prove $\mathbb{P}Z \exp(i\boldsymbol{\theta} \cdot \mathbf{Y}) = 0$ for all $\boldsymbol{\theta}$ in $\mathbb{R}^k$.
- Work with stochastic integral notation. Show that $\boldsymbol{\theta} \cdot \mathbf{Y} = H \bullet B_1$, where $H := \sum_{j=0}^{k-1} \theta_j ((t_j, t_{j+1}]]$.
- Show that $H \bullet B$ has a deterministic quadratic variation process, $A_t := [H \bullet B, H \bullet B]_t = \int_0^t H^2(s)\, ds$.
- Use the results from Section 2 to show that
$$W_1 = 1 + i(WH) \bullet B_1 \qquad \text{where } W_t := \exp(iH \bullet B_t + {}^1\!/\!_2 A_t).$$
- Deduce that
$$\exp(A_1/2)\mathbb{P}Z \exp(i\boldsymbol{\theta} \cdot \mathbf{Y}) = 0.$$

□ • Are we done?

<6> **Corollary.** *For each local martingale* M *adapted to the Brownian filtration there exists an* H *in* $\text{loc}\mathcal{L}^2(\mu)$ *such that* $M_t = M_0 + (H \bullet B)_t$ *for* $0 \le t \le 1$.

Proof. Without loss of generality, suppose $M_0 = 0$. Define stopping times $\tau_k := 1 \wedge \inf\{t : |M_t| \ge k\}$.

- Why does $M_{\wedge \tau_k}$ belong to $\mathcal{M}_0^2[0, 1]$?
- For each k, explain why there exists an $H_k \in \mathcal{L}_2(\mu)$ such that
$$M_{t \wedge \tau_k} = \big(H_k((0, \tau_k]]\big) \bullet B_t \qquad \text{for } 0 \le t \le 1.$$
- Deduce that $(H_k((0, \tau_k]]) \bullet B_1 = (H_{k+1}((0, \tau_k]] \bullet B_1$ almost surely.
- Deduce that $H_k((0, \tau_k]] - H_{k+1}((0, \tau_k]] = 0$ almost everywhere $[\mu]$.
- Show that the H_k processes can be pasted together to create an H in $\text{loc}\mathcal{L}^2(\mu)$ for which $M_t = H \bullet B_t$ almost surely. □

Am I just repeateing the construction for the $\text{loc}\mathcal{M}_0^2[0, 1]$ stochastic integral?

REMARK. Should I extend to general $\mathcal{F}_1$-measurable random variables, perhaps using the method of Dudley (1977), getting a representation $Y_0 + H \bullet B_1$ with $H \in \text{loc}\mathcal{H}_{\text{Bdd}}$.

5. Problems

[1] Suppose $Z \in \mathcal{FV}_0 \cap \text{loc}\mathcal{M}_0^2(\mathbb{R}^+)$ and Z has continuous sample paths. Show that $Z_t = 0$ almost surely, for each t. Hint: Use the fact that $[Z, Z] = 0$ to deduce that $Z^2 = 2Z \bullet Z \in \text{loc}\mathcal{M}_0^2(\mathbb{R}^+)$. Find a sequence of stopping times $\tau_k \uparrow \infty$ for which $\mathbb{P}Z_{t\wedge\tau_k}^2 = 0$ for each t.

[2] Suppose $M \in \text{loc}\mathcal{M}_0^2(\mathbb{R}^+)$ has continuous sample paths. Suppose $A \in \mathcal{FV}_0$ also has continuous paths and $M^2 - A \in \text{loc}\mathcal{M}_0^2(\mathbb{R}^+)$. Deduce that $A = [M, M]$. Hint: Apply Problem [1] to $[M, M] - A$.

[3] Let X be an integrable random variable, and $\mathbf{Y} = (Y_1, \dots, Y_k)$ be a vector of random variables such that $\mathbb{P}X \exp(i\boldsymbol{\theta} \cdot \mathbf{Y}) = 0$ for all $\boldsymbol{\theta} = (\theta_1, \dots, \theta_k)$ in $\mathbb{R}^k$. Show that $\mathbb{P}(Xg(\mathbf{Y})) = 0$ for all bounded, measurable g. Hint: Let $\mu^{\pm}$ be the measures with densities $X^{\pm}$ with respect to $\mathbb{P}$. Show that $\mathbf{Y}$ has the same Fourier transform, and hence the same distribution, under both μ^+ and μ^-. That is, $\mu^+ g(\mathbf{Y}) = \mu^- g(\mathbf{Y})$.

[4] Suppose $\{X_t : 0 \le t \le 1\}$ is a cadlag martingale with respect to the Brownian filtration. Remember that X_1 can be expressed as $f(B)$ for some $\mathcal{C}\backslash\mathcal{B}(\mathbb{R})$-measurable functional f on $C[0, 1]$. The functional is $\mathbb{W}$-integrable.

(i) If f is a continuous (for sup-norm distance) functional on $C[0, 1]$, use the representation $X_t = \mathbb{P}_t f(B) = \mathbb{W}^x f(K_t B + S_t x)$ almost surely to show that X has continuous sample paths (almost surely?).

(ii) For a general $\mathbb{W}$-integrable functional, show that there exists a sequence of continuous functionals $\{f_n\}$ for which $\mathbb{W}|f - f_n| \le 4^{-n}$.

(iii) Let M_n be a version of the martingale $\mathbb{P}_t f_n(B)$ with continuous sample paths. Show that $|M_n(t) - X(t)|$ is a uniformly integrable submartingale with cadlag sample paths.

(iv) Define stopping times $\tau_n := 1 \wedge \min\{t : |M_n(t) - X(t)| \ge 2^{-n}\}$. Show that

$$\begin{aligned}\mathbb{P}\{\sup_t |M_n(t) - X(t)| > 2^{-n}\} &\le 2^n \mathbb{P}|M_n(\tau_n) - X(\tau_n)| \\ &\le 2^n \mathbb{P}|M_n(1) - X(1)| = 2^n \mathbb{P}|f_n(B) - f(B)| \\ &\le 2^{-n}\end{aligned}$$

(v) Deduce that $\sum_n \mathbb{P}\{\sup_t |M_n(t) - X(t)| > 2^{-n}\} < \infty$ and hence $\sup_t |M_n(t) - X(t)| \to 0$ almost surely.

(vi) Conclude that almost all sample paths of X are continuous.

(vii) Extend the argument to the case of a local martingale. Hint: If $M_{\wedge\tau_k}$ has (almost all) continuous paths for each k, and if $\tau_k \uparrow \infty$, what do you know about almost all paths of M?

[5] Let X and Y be independent Brownian Motions.

(i) Show that both $(X + Y)/\sqrt{2}$ and $(X - Y)/\sqrt{2}$ are also Brownian Motions.

(ii) Deduce that $[X, Y] = 0$.

The next problem presents the standard example of a uniformly integrable local martingale that is not of class [D].

Better to start at origin, and work with distance to $\mathbf{u}$?

[6] Let $\mathbf{B} = (1 + X, Y, Z)$ be a three-dimensional Brownian Motion started from $\mathbf{u} = (1, 0, 0)$. (The three processes X, Y, and Z are independent Brownian Motions started from zero.) Write $\mathbf{x}$ for (x, y, z). Define $f(\mathbf{x}) = 1/\|\mathbf{x}\|$ on $\mathbb{R}^3\backslash\{0\}$. Define a process $M(t) = 1/\|\mathbf{B}(t)\|$.

(i) Use the Multiprocess Itô Formula to show that $M \in \text{loc}\mathcal{M}^2(\mathbb{R}^+)$. Hint: Show that on the open region $\mathbb{R}^3\backslash\{0\}$ the function f is harmonic:

$$\frac{\partial^2 f}{\partial^2 x} + \frac{\partial^2 f}{\partial^2 y} + \frac{\partial^2 f}{\partial^2 z} = 0.$$

(ii) Deduce that M is a positive supermartingale.

(iii) Let $\tau_k = \inf\{t : \|\mathbf{B}(t)\| \le 1/k\}$. Show that $M_{\wedge\tau_k} \in \mathcal{M}^2(\mathbb{R}^+)$.

(iv) Show that $C_0 := \int\{\|\mathbf{x}\| \le 1/2\}\|\mathbf{x}\|^{-2}\, d\mathbf{x} < \infty$.

(v) Show that $\mathbb{P}M(t)^2 \le C_0 \exp(-(8t)^{-1})t^{-3/2} + \mathbb{P}\left(8 \wedge \|\mathbf{B}(t)\|^{-2}\right)$.

(vi) Show that $\|\mathbf{B}(t)\|^2 \xrightarrow{\mathbb{P}} \infty$ as $t \to \infty$.

(vii) Deduce that $\sup_t \mathbb{P}M(t)^2 < \infty$ and $\mathbb{P}M(t) \to 0$ as $t \to \infty$.

(viii) Deduce that M is not a martingale, and hence M is not in class [D].

References

Dudley, R. M. (1977), 'Wiener functionals as Itô integrals', *Annals of Probability* **5**, 140–141.

Protter, P. (1990), *Stochastic Integration and Differential Equations*, Springer, New York.

PROJECT 10

1. Change of measure for Brownian motion

Let $\{B_t : 0 \le t \le 1\}$ be a Brownian motion with respect to a (standard) filtration $\{\mathcal{F}_t\}$ on $(\Omega, \mathcal{F}, \mathbb{P})$. Write $\mathcal{U}$ for its quadratic variation process, $\mathcal{U}_t = t$.

For each $\alpha \in \mathbb{R}$, the process

$$q_t = \exp\left(\alpha B_t - \tfrac{1}{2}\alpha^2 t\right) \qquad \text{for } 0 \le t \le 1$$

is a nonnegative martingale, with $\mathbb{P}q_t = \mathbb{P}q_0 = 1$. Define a new probability measure $\mathbb{Q}_\alpha$ on $\mathcal{F}_1$ by specifying q_1 to be its density with respect to $\mathbb{P}$. That is,

$$\mathbb{Q}_\alpha X = \mathbb{P}(X q_1)$$

at least for all bounded random variables X.

- Show that $\mathbb{Q}_\alpha$ is equivalent to $\mathbb{P}$, in the sense that both measures have the same collection $\mathcal{N}$ of negligible sets.
- Show that $\mathbb{Q}_\alpha X = \mathbb{P}(X q_t)$ if X is $\mathcal{F}_t$-measurable. Explain why q_t is a Radon-Nikodym density for $\mathbb{Q}_\alpha$ with respect to $\mathbb{P}$ when both measures are restricted to $\mathcal{F}_t$.
- For fixed s and $t = s + \delta$, a fixed F in $\mathcal{F}_s$, and a bounded measurable f, show that

$$\begin{aligned}\mathbb{Q}_\alpha F f(B_t - B_s) &= \mathbb{P}(F q_s)\,\mathbb{P}\left(f(B_t - B_s)\exp\left(\alpha(B_t - B_s) - \tfrac{1}{2}\alpha^2\delta\right)\right)\\ &= \mathbb{Q}_\alpha F \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\delta}} f(z)\exp\left(-\tfrac{1}{2}(z - \alpha\delta)^2/\delta\right)\,dz\end{aligned}$$

- Deduce that, under $\mathbb{Q}_\alpha$, the process $B_t - \alpha t$ is a standard Brownian motion.

2. The Black-Scholes formula

Stock prices (in units so that $S_0 \equiv 1$) are sometimes modeled by a continuous process driven by a Brownian motion, B, on $[0, 1]$;

$$\begin{aligned}S_t &= \exp((\mu - {}^1\!/\!_2\sigma^2)t + \sigma B_t) \qquad \text{for } 0 \le t \le 1\\ &= \exp(\sigma \widetilde{B}_t - {}^1\!/\!_2\sigma^2 t) \qquad \text{where } \widetilde{B}_t = B_t + (\mu/\sigma)t\end{aligned}$$

for constants $\sigma > 0$ (assumed known) and μ (unknown). That is,

$$S_t = \psi(B_t, \mathcal{U}_t) \qquad \text{where } \psi(x, y) = \exp(\sigma x + (\mu - {}^1\!/\!_2\sigma^2)y).$$

I am ignoring inflation. cf. expression of value of stock as a multiple of a bond price.

Suppose $Y = f(S)$, with f a $\mathcal{C}$-measurable functional on $C[0, 1]$. How much should one pay at time 0 in order to receive the amount Y at time 1?

- Use the Itô formula to show that

<1> $$S_t = 1 + \sigma S \bullet B_t + \mu S \bullet \mathcal{U}_t,$$

In more traditional notation,

$$dS_t = \sigma S_t\, dB_t + \mu S_t\, dt, \qquad \text{or} \qquad \frac{dS_t}{S_t} = \sigma\, dB_t + \mu\, dt.$$

Roughly speaking, the relative increments of S behave like the increments of a Brownian motion with drift μ. The process $\sigma S \bullet B$ is the $\text{loc}\mathcal{M}_0^2[0, 1]$ part of the semimartingale decomposition of S.

- Similarly, show that $S_t = 1 + \sigma S \bullet \widetilde{B}_t$.
- Show that Y can be written as a $\mathcal{C}$-measurable functional of the $\widetilde{B}$ sample path.
- Temporarily suppose that $\mu = 0$, so that $\widetilde{B}$ is a standard Brownian motion.
  (i) Use stochastic calculus to show that
  $$\widetilde{B} = \frac{1}{\sigma S} \bullet S$$
  Hint: What do you know about the increments of the process that takes a constant value?
  (ii) Suppose $\mathbb{P}Y^2 < \infty$. Invoke results from Project 9 to show that there exists a predictable H such that
  $$Y = \mathbb{P}Y + H \bullet \widetilde{B}_1 = \mathbb{P}Y + K \bullet S_1 \qquad \text{where } K := \frac{H}{\sigma S}. \tag{2}$$
  (iii) Interpret the last equality as an assertion that there exists an (idealized?) hedging stategy that returns $Y - \mathbb{P}Y$. Deduce that the arbitrage price for Y equals $\mathbb{P}Y$ in the special case where $\mu = 0$.
- Now consider the case where μ is unknown, possibly nonzero. Let $\mathbb{Q}_\alpha$ be the probability measure with density $\exp(\alpha B_1 - \frac{1}{2}\alpha^2)$ with respect to $\mathbb{P}$, where $\alpha = -\mu/\sigma$. Show that $\widetilde{B}$ is a standard Brownian motion under $\mathbb{Q}_\alpha$.
- Assume that $\mathbb{Q}_\alpha Y^2 < \infty$. Show that there exists some predictable process K_α (in some apppropriate $\mathcal{L}^2$ space) for which
  $$Y = \mathbb{Q}_\alpha Y + K_\alpha \bullet S_1 \qquad \text{almost surely } [\mathbb{Q}_\alpha].$$
- I believe that the threat of the trading scheme that delivers a return $K_\alpha \bullet S_1$ now forces $\mathbb{Q}_\alpha Y$ to be the amount one should pay at time 0 to receive the amount Y at time 1. What do you think? Should the fact that K_α seems to depend on the unknown μ invalidate the arbitrage argument?
- Suppose Y actually depends only on the stock price at time 1, that is, $Y = f_1(S_1)$ for some measurable function f_1. Show that
  $$\mathbb{Q}_\alpha Y = \mathbb{Q} f_1\left(\exp(\sigma W - \tfrac{1}{2}\sigma^2)\right) \qquad \text{where } W \sim N(0,1) \text{ under } \mathbb{Q}.$$
  Deduce that $\mathbb{Q}_\alpha Y$ does not depend on μ.
- Specialize even further, to the case where $f_1(x) = (x - C)^+$, for some constant C, to derive the famous Black-Scholes formula for the price of a European option.

3. Does K_α actually depend on μ?

I should reread Harrison & Pliska (1981).

As I type this Project late at night, I find myself in the embarrassing position of not really understanding how the question is handled for a general Y. However, when $Y = f_1(S_1)$ there is another approach that avoids the difficulty by constructing an explicit strategy via the solution to a partial differental equation.

Look for a smooth function $f(x,t)$ for which $f(x,1) = f_1(x)$ and
$$\sigma^2 x^2 f_{xx}(x,t) + f_t(x,t) = 0$$

- Use Itô to show that
  $$f(S_t, t) = f(1,0) + F_x \bullet S_t \qquad \text{almost surely } [\mathbb{P}].$$
- Show that, under $\mathbb{Q}_\alpha$, the stock price process is a martingale. Deduce that
  $$f(S_t, t) = \mathbb{Q}_\alpha\left(f(S_1, 1) \mid \mathcal{F}_t\right) = \mathbb{Q}_\alpha(Y \mid \mathcal{F}_t)$$

and, in particular, $f(1,0) = \mathbb{Q}_\alpha(Y \mid \mathcal{F}_0) = \mathbb{Q}_\alpha Y$.

I hope I can sort through my confusion before the lecture. I will reread the final section of Chung & Williams (1990).

4. Change of measure for semimartingales

The key fact about the change from $\mathbb{P}$ to an equivalent measure $\mathbb{Q}$ is the preservation of the semimartingale property. It is not at all an obvious fact. For suppose that X is a $\mathbb{P}$-semimartingale that has decomposition $X_0 + M + A$, where M is a locally square integrable $\mathbb{P}$-martingale and $A \in \mathcal{FV}_0$. Under $\mathbb{Q}$ the A process is still in $\mathcal{FV}_0$, but we will have to subtract another $\mathcal{FV}_0$ process A^* from M to make it a locally square integrable $\mathbb{Q}$-martingale, leading to the $\mathbb{Q}$-semimartingale decomposition $X = X_0 + (M - A^*) + (A + A^*)$.

To establish these facts in the general case I would need some theory about processes with jumps—things like the Doob-Meyer decomposition. Using only tools developed in the course, I can show you how to treat a special case.

Consider only a process $M \in \text{loc}\mathcal{M}_0^2([0,1],\mathbb{P})$ with continuous sample paths. Here I have added the $\mathbb{P}$ to emphasize that the martingale properties hold under the $\mathbb{P}$ distribution. Suppose that $\mathbb{P}$ and $\mathbb{Q}$ are equivalent measures, with $q_1 := d\mathbb{Q}/d\mathbb{P}$ and $d\mathbb{P}/d\mathbb{Q} = p_1 = 1/q_1$. Assume that the cadlag versions of the $\mathbb{P}$-martingale $q_t := \mathbb{P}(q_1 \mid \mathcal{F}_t)$ and the $\mathbb{Q}$-martingale $p_t := \mathbb{Q}(p_1 \mid \mathcal{F}_t)$ actually have continuous sample paths.

- Show that $p_\sigma = \mathbb{Q}(p_1 \mid \mathcal{F}_\sigma)$ for each $[0,1]$-valued stopping time σ.
- Explain why we can assume $p_t q_t \equiv 1$. More specifically, explain why p_t can be thought of as the density of $\mathbb{P}$ with respect to $\mathbb{Q}$ when both measures are restricted to $\mathcal{F}_t$.

Define

$$\tau_k := 1 \wedge \inf\{t : p_t \geq k \text{ or } p_t \leq 1/k\} \wedge \inf\{t : |M_t| \geq k\}.$$

Without loss of generality, we may also assume that $M_{\wedge\tau_k} \in \mathcal{M}_0^2([0,1],\mathbb{P})$.

- Show that $pM \in \text{loc}\mathcal{M}_0^2([0,1],\mathbb{Q})$. Hint: For $s < t$ and $F \in \mathcal{F}_s$ show that
$$\mathbb{Q}F\left(p_{t\wedge\tau_k} M_{t\wedge\tau_k} - p_{s\wedge\tau_k} M_{s\wedge\tau_k}\right) = \mathbb{Q}F\left(p_{t\wedge\tau_k} M_{t\wedge\tau_k} - p_{s\wedge\tau_k} M_{s\wedge\tau_k}\right)\{\tau_k > s\}$$
Argue that $F\{\tau_k > s\} \in \mathcal{F}_{s\wedge\tau_k}$ then deduce that the right-hand side of the last equality equals $\mathbb{P}F\left(M_{t\wedge\tau_k} - M_{s\wedge\tau_k}\right)\{\tau_k > s\} = 0$.
- Use the fact that q and M are both in $\text{loc}\mathcal{M}_0^2([0,1],\mathbb{P})$ to explain why the process $Y := qM - V$, where $V := [q, M] \in \mathcal{FV}_0$, is also in $\text{loc}\mathcal{M}_0^2([0,1],\mathbb{P})$. Hint: First explain why $Y_{t\wedge\tau_k} = q \bullet M_{t\wedge\tau_k} + M \bullet q_{t\wedge\tau_k}$.
- Explain why both Y and V have continuous sample paths.
- Explain why $pY \in \text{loc}\mathcal{M}_0^2([0,1],\mathbb{Q})$.
- Explain why $[p, V] \equiv 0$. Hint: V is a $\mathcal{FV}_0$ process with continuous sample paths.
- Deduce that $p_t V_t = p \bullet V_t + V \bullet p_t$.
- Deduce that $M - p \bullet V - V \bullet p \in \text{loc}\mathcal{M}_0^2([0,1],\mathbb{Q})$.

- Explain why $V \bullet p \in \text{loc}\mathcal{M}_0^2([0,1],\mathbb{Q})$. Hint: p is a $\mathbb{Q}$-martingale.
- Explain why $A := p \bullet V$ is in $\mathcal{FV}_0$.
- Conclude that $M - A \in \text{loc}\mathcal{M}_0^2([0,1],\mathbb{Q})$.

You should check that this recipe works for the Brownian motion example in Section 1.

5. Things I could show if I had more time

(Actually I would also need some facts about processes with jumps.)

(i) Every local-martingale is a semimartingale.

(ii) Suppose $\mathbb{P}$ and $\mathbb{Q}$ are equivalent probability measures. If X is a $\mathbb{P}$-semimartingale then it is also a $\mathbb{Q}$-semimartingale. Moreover, for $H \in \text{loc}\mathcal{H}_{\text{Bdd}}$, the stochastic integral $H \bullet X$ when calculated using the methods from Project 7 under $\mathbb{P}$ is the same as the stochastic integral when calculated under $\mathbb{Q}$. The last assertion can be proved using the characterization of the stochastic integral given in Project 7.

References

Chung, K. L. & Williams, R. J. (1990), *Introduction to Stochastic Integration*, Birkhäuser, Boston.

Harrison, J. M. & Pliska, S. R. (1981), 'Martingales and stochastic integrals in the theory of continuous trading', *Stochastic Processes and their Applications* **11**, 215–260.

Project 11

Notation: Write $\mathbb{P}_t(\ldots)$ for $\mathbb{P}(\ldots \mid \mathcal{F}_t)$ and $\text{var}_t(\ldots)$ for the corresponding conditional variance.

1. Diffusion heuristics

The rough idea of an Itô diffusion is: $\{X_t : t \in \mathbb{R}^+\}$ is adapted with continuous sample paths; and for small $\delta > 0$, with $\Delta X = X_{t+\delta} - X_t$,

$$\mathbb{P}_t\left(\Delta X\right) \approx \delta b(X_t) \tag{<1>}$$

$$\text{var}_t\left(\Delta X\right) \approx \delta \sigma^2(X_t) \tag{<2>}$$

where $b(\cdot)$ and $\sigma(\cdot)$ are deterministic functions. In what follows, both b and σ will be continuous functions.

Interpret <1> to mean that

$$\mathbb{P}_t(\Delta Z) \approx 0 \qquad \text{where } Z_t = X_t - \int_0^t b(X_s)\,ds.$$

More precisely, interpret <1> to mean that Z is a martingale with continuous sample paths and $Z_0 = 0$. Similarly, interpret <2> to mean $\mathbb{P}_t(\Delta Z)^2 \approx \delta\sigma^2(X_t)$, or

$$W_t := [Z, Z]_t - \int_0^t \sigma^2(X_s)\,ds \qquad \text{is a martingale.}$$

Note that W has continuous paths of finite variation. From the Problems to Project 9, we must have $W_t \equiv W_0 = 0$. That is, $[Z, Z]_t = \int_0^t \sigma^2(X_s)\,ds$.

Put another way, we could interpret <1> and <2> to mean that

$$X_t = x_0 + Z_t + b(X) \bullet \mathcal{U}_t \qquad \text{where } X_0 = x_0 \tag{<3>}$$

with Z a (local?) martingale for which $[Z, Z] = \sigma^2(X) \bullet \mathcal{U}$. Here, and subsequently, I am abusing notation by writing $b(X)$ for the process that takes the value $b(X_s)$ at time s, and so on.

Suppose there exist processes X and Z with the properties just described. If $\sigma(x) \neq 0$ for all x then $1/\sigma(X)$ is locally bounded and predictable. The process $B := (1/\sigma(X)) \bullet Z$ is a local martingale, with continuous sample paths, $B_0 = 0$, and

Note that $\sigma^2(X)$ is adapted and has continuous paths

$$[B, B] = (1/\sigma^2(X)) \bullet [Z, Z] = \mathcal{U}.$$

Compare with the argument in Stroock & Varadhan (1979, Section 4.5)

That is, by the Lévy characterization, B is a Brownian motion for which

$$X_t = x_0 + \sigma(X) \bullet B_t + b(X) \bullet \mathcal{U}_t \tag{<4>}$$

Many authors would write the last representation as

$$dX_t = \sigma(X_t)\,dB_t + b(X_t)\,dt \tag{<5>}$$

and call it a ***stochastic differential equation*** for X with initial condition $X_0 = x_0$.

If the representation <3> were valid, and if f were twice continuously differentiable, Itô's formula would give

$$\begin{aligned} f(X_t) &= f(x_0) + f'(X) \bullet (Z + b(X) \bullet \mathcal{U})_t + \tfrac{1}{2} f''(X) \bullet [Z, Z]_t \\ &= f(x_0) + f'(X) \bullet Z_t + \left(\tfrac{1}{2}\sigma^2(X) f''(X) + b(X) f'(X)\right) \bullet \mathcal{U}_t \end{aligned}$$

This representation would imply that

$$f(X_t) - \left(\tfrac{1}{2}\sigma(X)^2 f''(X) + b(X) f'(X)\right) \bullet \mathcal{U}_t \qquad \text{is a martingale} \tag{<6>}$$

for each suitably smooth f.

The question of whether an X satisfying <4> or <6>actually exists, and to what extent it is uniquely determined, is the subject of a huge literature. The small sampling that follows is based mostly on

(i) Stroock & Varadhan (1979, Chapters 4 and 5),

(ii) Durrett (1984, Chapter 9)

(iii) Chung & Williams (1990, Chapter 10).

2. Existence and uniqueness of a solution to a SDE

SDE = stochastic differential equation

Seek a solution for the SDE <5> with initial condition $X_0 \equiv x_0$, for a fixed $x_0 \in \mathbb{R}$. Suppose the functions b and σ satisfy the following conditions for some finite constant C:

<7>
$$\begin{cases} |b(x)| \le C, & |\sigma(x)| \le C \qquad \text{for all } x \\ |b(x) - b(y)| \le C|x-y|, & |\sigma(x) - \sigma(y)| \le C|x-y| \qquad \text{for all } x \text{ and } y \end{cases}$$

I am so very lazy to use the same constant for all the bounds.

Assume a standard Brownian motion B is given. Start by building the solution on a fixed interval $[0,T]$. Define $X^{(0)} \equiv x_0$ and, for $n \ge 0$,

$$X_t^{(n+1)} = x_0 + \sigma(X^{(n)}) \bullet B_t + b(X^{(n)}) \bullet \mathfrak{U}_t$$

Define

$$\Delta_{n+1}(t) := \mathbb{P}\sup\nolimits_{s\le t}|X_s^{(n+1)} - X_s^{(n)}|^2.$$

- Show that $\Delta_1(T) \le c_0 := 8C^2T + 2C^2T^2$, or something like that.
- For $n \ge 1$ show that

$$\begin{aligned}\Delta_{n+1}(T) &\le 2\mathbb{P}\sup\nolimits_{t\le T}|\sigma(X^{(n)}) \bullet B_t - \sigma(X^{(n-1)}) \bullet B_t|^2 \\ &\quad + 2\mathbb{P}\sup\nolimits_{t\le T}\left|\int_0^t b(X_s^{(n)}) - b(X^{(n-1)})\,ds\right|^2 \\ &\le 8\mathbb{P}|\sigma(X^{(n)}) \bullet B_T - \sigma(X^{(n-1)}) \bullet B_T|^2 \\ &\quad + 2T^2\mathbb{P}\left(\frac{1}{T}\int_0^T |b(X_s^{(n)}) - b(X_s^{(n-1)})|\,ds\right)^2 \\ &\le 8\int_0^T \mathbb{P}|\sigma(X_s^{(n)}) - \sigma(X_s^{(n-1)})|^2 \\ &\quad + 2T^2\mathbb{P}\left(\frac{1}{T}\int_0^T |b(X_s^{(n)}) - b(X_s^{(n-1)})|\,ds\right)^2 \\ &\le K_T\int_0^T \Delta_n(s)\,ds,\end{aligned}$$

  where K_T is a constant that depends on T.
- Strengthen the previous result to

$$\Delta_{n+1}(t) \le K_T\int_0^t \Delta_n(s)\,ds \qquad \text{for all } t \in [0,T].$$

- Show that

$$\begin{aligned}\Delta_{n+1}(T) &\le K_T^n \int \dots \int \{0 \le t_1 \le t_2 \le \dots \le t_n \le T\}\Delta_1(t_1)\,dt_1 dt_2 \dots dt_n \\ &\le c_0(TK_T)^n/n!\end{aligned}$$

- Deduce that

$$\mathbb{P}\sum\nolimits_{n\ge1}\sup\nolimits_{s\le T}|X_s^{(n+1)} - X_s^{(n)}| < \infty$$

- Deduce that there exists an adapted process $\{X_t : 0 \le t \le T\}$ with continuous sample paths, such that
$$\sup_{s\le T} |X_s^{(n)} - X_s| \to 0 \qquad \text{almost surely.}$$
- Deduce that
$$\sup_{s\le T} \left(|b(X_s^{(n)}) - b(X_s)| + |\sigma(X_s^{(n)}) - \sigma(X_s)|\right) \to 0 \qquad \text{almost surely.}$$
- Deduce that
$$|\sigma(X^{(n)}) \bullet B - \sigma(X) \bullet B| + |b(X^{(n)}) \bullet \mathcal{U} - b(X) \bullet \mathcal{U}| \xrightarrow{ucpc} 0$$
- Conclude that $\{X_t : 0 \le t \le T\}$ satisfies the SDE <5> with initial condition $X_0 \equiv x_0$.
- Suppose $\{Y_t : 0 \le t \le T\}$ is another solution to the SDE with the same initial condition. Define
$$\Delta(t) := \mathbb{P} \sup_{s\le t} |X_s - Y_s|^2.$$
Show that for some constants c_1 and κ, which might depend on T,
$$\Delta(T) \le \left(c_1 \kappa^n / n!\right) \Delta(T).$$
Deduce that $\Delta(T) = 0$ and hence
$$\mathbb{P}\{\omega : \exists t \le T \text{ with } X_t(\omega) \ne Y_t(\omega)\} = 0.$$
- Suppose $\{X_t : 0 \le t \le T_1\}$ and $\{Z_t : 0 \le t \le T_2\}$ are solutions to the SDE over different ranges, $[0, T_1]$ and $[0, T_2]$, with $X_0 = Z_0 = x_0$. Show that almost all paths $X(\cdot, \omega)$ and $Z(\cdot, \omega)$ agree on the interval $[0, T_1 \wedge T_2]$. Explain how this result enables us to find a unique solution (up to almost sure equivalence) on $\mathbb{R}^+$.

3. Dependence of the solution on B: strong and weak solutions of the SDE

The solution X constructed in Section 2 depends only on the Brownian motion. More precisely, we could choose $\{\mathcal{F}_t\}$ as the augmented Brownian filtration and have X adapted to that filtration.

- Try to make some sense of the last assertion. Perhaps you could argue inductively that each approximation $X^{(n)}$ is adapted to the augmented filtration. I would like to show that this means we can choose $X_t(\omega)$ as $f(B_{\wedge t}(\omega), t)$ for some suitably measurable function $f : C(\mathbb{R}^+) \times \mathbb{R}^+ \to \mathbb{R}$. Perhaps we could require $t \mapsto f(y, t)$ to be continuous for each fixed y.

The idea is that B can provide both the filtration and the process for the stochastic integral $\sigma(X) \bullet B$. I think this is what it means for X to be a ***strong solution*** of SDE. Clearly, if we start from a different Brownian motion then we get a different solution.

The distribution of X is a probability measure, $\mathbb{Q}_{x_0}$, on the cylinder sigma-field $\mathcal{C}$ of $C(\mathbb{R}^+)$. More formally, if we can regard f as a $\mathcal{C}\backslash\mathcal{C}$-measurable map from $C(\mathbb{R}^+)$ back into itself, then $\mathbb{Q}_{x_0}$ is the image of Wiener measure $\mathbb{W}$ under the map f.

I am a lttle unsure of these assertions, because I have not worked through the whole construction myself. I am relying on what I think Durret and Chung&Williams are asserting.

I think that for some SDE's it is possible to prove the existence of a $\mathbb{Q}_{x_0}$ on $\mathcal{C}$ under which the coordinate map defines a process with continuous paths started at x_0 for which the analog of property <6> holds. Slight refinements of the arguments in Section 1 then show how to construct a Brownian motion B for which <4> holds.

For a famous example where there exists a (nonunique) weak solution but no strong solution see Chung & Williams (1990, Secton 10.4).

4. Relaxation of assumptions on b and σ

Localization arguments allow us to relax the conditions <7> on the functions $b(\cdot)$ and $\sigma(\cdot)$ to existence of constants C_r for each $R > 0$ such that

<8> $$\max\big(|b(x) - b(y)|, |\sigma(x) - \sigma(y)|\big) \le C_R|x - y| \qquad \text{if } \max(|x|, |y|) \le R.$$

Most authors seem also to require a growth condition,

$$\max\big(|b(x)|, |\sigma(x)|\big) = O(|x|) \qquad \text{as } |x| \to \infty.$$

Frankly, I do not really understand why the growth condition is needed.

It seems to me that assumption <8> implies existence of finite constants K_R for which

$$|b(x)| + |\sigma(x)| \le K_R \qquad \text{when } |x| \le R.$$

Define

$$b_R(x) := \max(-K_R, \min(b(x), K_R))$$
$$\sigma_R(x) := \max(-K_R, \min(\sigma(x), K_R))$$

An analog of <7> holds for b_R and σ_R. There exists continuous adapted processes for which

$$X_t^{(R)} = x_0 + \sigma_R(X^{(R)}) \bullet B_t + b_R(X^{(R)}) \bullet \mathcal{U}_t$$

Define $\tau_R := \inf\{t : |X_t^{(R)}| \ge R\}$. I think that

$$X_{t\wedge\tau_R}^{(R)} = x_0 + \sigma(X^{(R)}) \bullet B_{t\wedge\tau_R} + b(X^{(R)}) \bullet \mathcal{U}_{t\wedge\tau_R}$$

It should be possible to paste together the solutions $X^{(R)}$ for an increasing sequence of R values, invoking the uniqueness theorem from Section 2 to show that $X^{(2R)}$ agrees with $X^{(R)}$ at least until $|X^{(2R)}| \ge R$. If the corresponding stopping times τ_R were to increase to infinity as $R \uparrow \infty$ then we would get a solution to the original SDE. I think this is where the growth condition is needed.

I need to read the last part of Chung & Williams (1990, Secton 10.2) more carefully.

5. Examples

We should try to establish existence and uniqueness of the solutions to two simple SDE's:

(i) (geometric Brownian motion) Using the Itô formula, you showed in Project 10 that

$$X_t = \exp\big(\sigma B_t + (\mu - \tfrac{1}{2}\sigma^2)t\big)$$

is a solution to the equation $X_t = 1 + \sigma X \bullet B_t + \mu X \bullet \mathcal{U}_t$. Is it the only solution?

(ii) (Ornstein-Uhlenbeck process) By the Itô formula, the process

$$X_t = e^{-\alpha t}\big(x_0 + E \bullet B_t\big) \qquad \text{where } E_s := e^{\alpha s}$$

is a solution to the SDE $dX_t = -\alpha X_t + dB_t$ with $X_0 = x_0$, that is,

$$X_t = x_0 + B_t - \alpha X \bullet \mathcal{U}_t$$

Again, is it the only solution? Could we establish both existence and uniqueness of a (strong) solution by appeal to the general theory?

References

Chung, K. L. & Williams, R. J. (1990), *Introduction to Stochastic Integration*, Birkhäuser, Boston.

Durrett, R. (1984), *Brownian Motion and Martingales in Analysis*, Wadsworth, Belmont CA.

Stroock, D. W. & Varadhan, S. R. S. (1979), *Multidimensional Diffusion Processes*, Springer, New York.

Doléans measures

Notation

- $\mathcal{T}_1$ = the set of all [0, 1]-valued stopping times
- for $\sigma, \tau \in \mathcal{T}_1$,

$$((\sigma, \tau]] := \{(t, \omega) \in (0, 1] \times \Omega : \sigma(\omega) < t \leq \tau(\omega)\}$$
$$[[\sigma, \tau]] := \{(t, \omega) \in (0, 1] \times \Omega : \sigma(\omega) \leq t \leq \tau(\omega)\}$$

and so on.

1. Introduction

The construction (Project 4) of the isometric stochastic integral $H \bullet M$ with respect to a martingale $M \in \mathcal{M}^2[0, 1]$, at least for bounded, predictable H, depended on the existence of the Doléans measure μ on the predictable sigma-field $\mathcal{P}$ on $(0, 1] \times \Omega$. To make the map $H \mapsto H \bullet M_1$ an isometry between $\mathcal{H}_{\text{simple}}$ and a subset of $\mathcal{L}^2(\Omega, \mathcal{F}_1, \mathbb{P})$ we needed

<1> $$\mu(a, b] \times F = \mathbb{P}\{\omega \in F\}(M_b - M_a)^2 \qquad \text{for all } 0 \leq a < b \leq 1 \text{ and } F \in \mathcal{F}_a.$$

⇛ This property characterizes the measure μ because the collection of all predictable sets of the form $(a, b] \times F$ is $\cap f$-stable and it generates $\mathcal{P}$.

The sigma-field $\mathcal{P}$ is also generated by the set of all stochastic intervals $((0, \tau]]$ for $\tau \in \mathcal{T}_1$. The Doléans measure is also characterized by the property ⇛

$$\mu((0, \tau]] = \mathbb{P}(M_\tau - M_0)^2 \qquad \text{for all } \tau \in \mathcal{T}_1.$$

Notice that μ depends on M only through the submartingale $S_t := \big(M_t - M_0\big)^2$:

$$\mathbb{P}F(M_b - M_a)^2 = \mathbb{P}F(S_b - S_a) \qquad \text{for } F \in \mathcal{F}_a.$$

In fact, analogous measures can be defined for a large class of submartingales.

<2> **Definition.** *Let $\{S_t : 0 \leq t \leq 1\}$ be a cadlag submartingale. Say that a finite (countably-additive) measure μ_S, defined on the predictable sigma-field of $(0, 1] \times \Omega$, is the* ***Doléans measure*** *for S if $\mu((0, \tau]] = \mathbb{P}\left(S_\tau - S_0\right)$ for every τ in $\mathcal{T}_1$.*

Remark.

If μ_S exists then $S_\tau - S_0$ must be integrable for every τ in $\mathcal{T}_1$. Note also that the definition is not affected if we replace S_t by $S_t - S_0$. Thus there is no loss of generality in assuming that $S_0 \equiv 0$.

I mentioned explicitly that μ_S must be countably-additive to draw attention to a subtle requirement on S for μ_S to exist, known somewhat cryptically as property [D]:

[D] $$\{S_\tau : \tau \in \mathcal{T}_1\} \text{ is uniformly integrable.}$$

<3> **Example.** If $M \in \mathcal{M}^2[0,1]$ then the submartingale $S_t = (M_t - M_0)^2$ has property $[D]$. Indeed, we know from Project 2 that $M_\tau = \mathbb{P}(M_1 \mid \mathcal{F}_\tau)$ for all $\tau \in \mathcal{T}_1$. Thus

$$0 \le \left(M_\tau - M_0\right)^2 \le \mathbb{P}\left(\left(M_1 - M_0\right)^2 \mid \mathcal{F}_\tau\right) \qquad \text{for each } \tau \text{ in } \mathcal{T}_1.$$

$\Rrightarrow$ Use the fact that $\{\mathbb{P}(\xi \mid \mathcal{G}) : \mathcal{G} \subseteq \mathcal{F}\}$ is uniformly integrable for each integrable random variable ξ to complete the argument. □

<4> **Example.** Let $\{B_t : 0 \le t \le 1\}$ be a standard Brownian motion. The submartingale $S_t := B_t^2$ has a very simple Doléans measure, characterized by

$$\mu_S(a,b] \otimes F = \mathbb{P}F\left(B_b^2 - B_a^2\right) = \mathbb{P}F(b-a) \qquad \text{for } F \in \mathcal{F}_a.$$

That is, $\mu_S = \mathfrak{m} \otimes \mathbb{P}$, with $\mathfrak{m}$ equal to Lebesgue measure on $\mathcal{B}[0,1]$. Of course μ_S has a further extension to the product sigma-field $\mathcal{B}[0,1] \otimes \mathcal{F}$.

A Poisson process $\{N_t : 0 \le t \le 1\}$ with intensity 1 shares with Brownian motion the independent increment property, but the increment $N_t - N_s$ has a Poisson$(t-s)$ distribution. The sample paths are constant, except for jumps of size 1 corresponding to points of the process. The process $\{N_t : 0 \le t \le 1\}$ is a submartingale with respect to its natural filtration, with Doléans measure $\mathfrak{m} \otimes \mathbb{P}$, the $\Rrightarrow$ same as the square of Brownian motion.

Clearly the Doléans measure does not uniquely determine the submartingale: both squared Brownian motion and the Poisson process have Doléans measure $\mathfrak{m} \otimes \mathbb{P}$. But the only square integrable martingale M with continuous sample paths and Doléans measure $\mu_M = \mathfrak{m} \otimes \mathbb{P}$ is Brownian motion: if $F \in \mathcal{F}_s$ and $s < t$ then $\mathbb{P}F(M_t^2 - M_s^2) = \mu_M(s,t] \otimes F = (t-s)\mathbb{P}F$, from which it follows that $M_t^2 - t$ is a martingale with respect to $\{\mathcal{F}_t\}$. It follows from Lévy's characterization that M is a Brownian motion. □

Problem [2] shows that if there exists a (countably-additive) Doléans measure μ_S then S has property [D]. The proof of the converse assertion is the main subject of this handout.

<5> **Theorem.** *Every cadlag submartingale $\{S_t : 0 \le t \le 1\}$ with property [D] has a countably additive Doléans measure μ_S.*

There are several ways to prove this assertion For example:

(i) Invoke an approximation by compact sets (Métivier 1982, Chapter 3).

(ii) For the square of a continuous $\mathcal{M}_0^2[0,1]$-martingale M, prove directly the existence of an increasing process A (the quadratic variation process) for which $M_t^2 - A_t$ is a martingale (Chung & Williams 1990, Section 4.4).

(iii) Do something very general, as in Dellacherie & Meyer (1982, §7.1).

I will present a different method, based on the identification of measures on $\mathcal{P}$ with a certain kind of linear functional defined on the vector space $\mathcal{H}_{\text{BddLip}}$ of all adapted, continuous processes H on $[0,1] \times \Omega$ for which there exists a finite constant C_H such that

(i) $|H(t,\omega)| \le C_H$ for all (t,ω).

(ii) $|H(s,\omega) - H(t,\omega)| \le C_H|t-s|$ for all s, t, and ω.

⇛ It is easy to show that $\mathcal{H}_{\text{BddLip}}$ generates a sigma-field $\mathcal{P}_0$ on $[0,1]\times\Omega$ for which

$$\mathcal{P} = \{D \cap \big((0,1]\times\Omega\big) : D \in \mathcal{P}_0\}.$$

As a consequence of the Theorem stated in Section 3, an increasing linear functional $\mu : \mathcal{H}_{\text{BddLip}} \to \mathbb{R}$ is defined by the integral with respect to a finite, countably additive measure on $\mathcal{P}_0$ if and only if it is ***sigma-smooth at*** 0, that is,

$$\mu(h_n) \downarrow 0 \qquad \text{for each } \{h_n : n \in \mathbb{N}\} \subseteq \mathcal{H}_{\text{BddLip}} \text{ with } h_n \downarrow 0 \text{ pointwise.}$$

If $\mu\big(\{0\}\times\Omega\big) = 0$ then μ can also be thought of as a measure on $\mathcal{P}$.

2. The Doléans measure as a linear functional

Let $\{S_t : 0 \le t \le 1\}$ be a cadlag submartingale with respect to a standard filtration $\{\mathcal{F}_t : 0 \le t \le 1\}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Suppose S has property [D]. To prove Theorem <5> we need to construct an increasing linear functional on $\mathcal{H}^+_{\text{BddLip}}$ that is sigma-smooth at 0.

Without loss of generality assume $S_0 \equiv 0$.

Construct μ as a limit of simpler increasing linear functionals on $\mathcal{H}^+_{\text{BddLip}}$. For each n in $\mathbb{N}$ and $i = 0, 1, \dots, 2^n$ define $t_{i,n} := i/2^n$ and $\Delta_{i,n} := S(t_{i+1,n}) - S(t_{i,n})$ and write $\mathbb{P}_{i,n}(\cdots)$ for expectations conditional on $\mathcal{F}(t_{i,n})$. Note that $\mathbb{P}_{i,n}\Delta_{i,n} \ge 0$ almost surely, by the submartingale property.

For each H in $\mathcal{H}^+_{\text{BddLip}}$, define linear functionals

$$\mu_n H := \sum_{0\le i<2^n} \mathbb{P}\big(H(t_{i,n})\Delta_{i,n}\big) = \sum_{0\le i<2^n} \mathbb{P}\big(H(t_{i,n})\mathbb{P}_{i,n}\Delta_{i,n}\big).$$

The second form ensures that μ_n is an increasing functional on $\mathcal{H}^+_{\text{BddLip}}$.

Existence of the limit

To prove that $\mu H := \lim_{n\to\infty} \mu_n H$ exists for each $H \in \mathcal{H}^+$, I will show that the sequence $\{\mu_n H : n \in \mathbb{N}\}$ is Cauchy. Fix n and m with $n < m$. Define

$$J_i = \{j : t_{i,n} \le t_{j,m} < t_{i+1,n}\}.$$

Then

$$\begin{aligned}
&\Big|\Big(\sum\nolimits_{j\in J_i} \mathbb{P}H(t_{j,m})\Delta_{j,m}\Big) - \mathbb{P}H(t_{i,n})\Delta_{i,n}\Big| \\
&\qquad = \Big|\sum\nolimits_{j\in J_i} \mathbb{P}\big(H(t_{j,m}) - H(t_{i,n})\big)\Delta_{j,m}\Big| \\
&\qquad \le \sum\nolimits_{j\in J_i} \mathbb{P}\big(|H(t_{j,m}) - H(t_{i,n})|\mathbb{P}_{j,m}\Delta_{j,m}\big) \\
&\qquad \le \sum\nolimits_{j\in J_i} C_H 2^{-n}\mathbb{P}\Delta_{j,m} \\
&\qquad = C_H 2^{-n} \sum\nolimits_{j\in J_i} \mathbb{P}\Delta_{i,n}
\end{aligned}$$

Sum over i to deduce that $|\mu_m H - \mu_n H| \le C_H 2^{-n}\mathbb{P}S_1$, which tends to zero as n tends to infinity.

A useful upper bound

The functional μ inherits linearity and the increasing property from the $\{\mu_n\}$. For each fixed $\epsilon > 0$, property [D] will give an upper bound for μH in terms of the stopping time

<6> $$\tau(H,\epsilon) := \inf\{t : H(t,\omega) \geq \epsilon\} \wedge 1.$$

Temporarily write τ_n for the discretized stopping time obtained by rounding $\tau(H,\epsilon)$ up to the next integer multiple of 2^{-n}. Then

$$\begin{aligned}\mu_n H &\leq \sum\nolimits_{0\leq i<2^n} \mathbb{P}\left(\epsilon\{t_{i,n} < \tau_n\} + C_H\{t_{i,n} \geq \tau_n\}\right)\mathbb{P}_{i,n}\Delta_{i,n}\\ &\leq \epsilon \sum\nolimits_{0\leq i<2^n} \mathbb{P}\Delta_{i,n} + C_H \sum\nolimits_{0\leq i<2^n} \mathbb{P}\{t_{i,n} \geq \tau_n\}\Delta_{i,n}\\ &\leq \epsilon\mathbb{P}S_1 + C_H\mathbb{P}\left(S_1 - S_{\tau_n}\right).\end{aligned}$$

Let n tend to infinity. Uniform integrability of the sequence $\{S_{\tau_n}\}$ together with right-continuity of the sample paths of S lets us deduce in the limit that

<7> $$\mu H \leq \epsilon\mathbb{P}S_1 + C_H\mathbb{P}\left(S_1 - S_{\tau(H,\epsilon)}\right).$$

Sigma-smoothness

Now suppose $\{H_k : k \in \mathbb{N}\}$ is a sequence from $\mathcal{H}^+_{\text{BddLip}}$ for which $1 \geq H_k \downarrow 0$ pointwise. For a fixed $\epsilon > 0$, temporarily write σ_k for $\tau(H_k,\epsilon)$. By compactness of $[0,1]$, the pointwise convergence of the continuous functions, $H_k(\cdot,\omega) \downarrow 0$, is actually uniform. For each ω, the sequence $\{\sigma_k(\omega)\}$ not only increases to 1, it actually achieves the value 1 at some finite k (depending on ω). Uniform integrability of $\{S_{\sigma_k} : k \in \mathbb{N}\}$ and the analog of <7> for each H_k then give

$$\mu H_k \leq \epsilon\mathbb{P}S_1 + C_H\mathbb{P}\left(S_1 - S_{\sigma_k}\right) \to \epsilon\mathbb{P}S_1 \qquad \text{as } k \to \infty.$$

The sigma-smoothness of μ follows. The functional corresponds to the integral with respect to a finite measure on $\mathcal{P}$, with total mass $\mu[[0,1]] = \lim_n \mu_n[[0,1]] = \mathbb{P}S_1$.

Identification as a Doléans measure

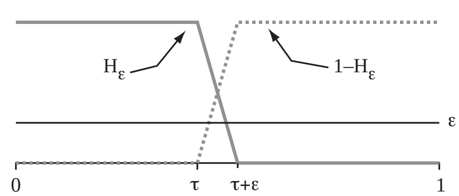
It remains to prove that

(a) $\mu\{0\} \times \Omega = 0$

(b) $\mu[[0,\tau]] = \mathbb{P}S_\tau$ for $\tau \in \mathcal{T}_1$.

Consider first the proof of (b). For given $\epsilon > 0$, approximate $[[0,\tau]]$ by the continuous process

$$H_\epsilon(t,\omega) = \min\left(1, (\tau(\omega) + \epsilon - t)^+/\epsilon\right) \qquad \text{for } 0 \leq t \leq 1.$$



It is adapted because $\{H_\epsilon(t,\omega) \le c\} = \{\tau \le t - \epsilon(1-c)\} \in \mathcal{F}_t$, for each fixed t and constant $0 \le c < 1$. It belongs to $\mathcal{H}^+_{\text{BddLip}}$, with $C_{H_\epsilon} = 1/\epsilon$, and $[[0,\tau]] \le H_\epsilon \le [[0,\tau+\epsilon]]$.

When $2^{-n} < \epsilon$, direct calculation shows that $\mu_n H_\epsilon \le \mathbb{P}S_{\tau+2\epsilon}$, which in the limit implies $\mu H_\epsilon \le \mathbb{P}S_{\tau+2\epsilon}$. By Dominated Convergence,

$$\mu[[0,\tau]] = \lim_{\epsilon\to 0} \mu H_\epsilon \le \mathbb{P}S_\tau.$$

Inequality <7> applied with $\tau_\epsilon := \tau(1-H_\epsilon, \epsilon)$ gives

$$\mu((\tau+\epsilon, 1]] \le \mu\left(1 - H_\epsilon\right) \le \epsilon\mathbb{P}S_1 + \mathbb{P}\left(S_1 - S_{\tau_\epsilon}\right),$$

which, in the limit as ϵ tends to zero, implies $\mu(\tau,1]] \le \mathbb{P}\left(S_1 - S_\tau\right)$. From the fact that

$$\mathbb{P}S_1 = \mu[[0,1]] = \mu[[0,\tau]] + \mu((\tau,1]] \le \mathbb{P}S_\tau + \mathbb{P}\left(S_1 - S_\tau\right),$$

conclude that $\mu[[0,\tau]] = \mathbb{P}S_\tau$.

□ Specialize to the case $\tau \equiv 0$ to get (a).

3. Measures as linear functionals

The following material on the Daniell construction of integrals is taken almost verbatim from Pollard (2001, Appendix A), where proofs are given.

<8> **Definition.** *Call a class $\mathcal{H}^+$ of nonnegative real functions on a set $\mathcal{X}$ a* ***lattice cone*** *if it has the following properties. If h, h_1 and h_2 belong to $\mathcal{H}^+$, and α_1 and α_2 are nonegative real numbers, then:*

(H1) $\alpha_1 h_1 + \alpha_2 h_2$ belongs to $\mathcal{H}^+$;

(H2) $h_1 \backslash h_2 := (h_1 - h_2)^+$ belongs to $\mathcal{H}^+$;

(H3) the pointwise minimum $h_1 \wedge h_2$ and maximum $h_1 \vee h_2$ belong to $\mathcal{H}^+$;

(H4) $h \wedge 1$ belongs to $\mathcal{H}^+$.

For a lattice cone $\mathcal{H}^+$, let $\mathcal{K}_0$ denote the class of all sets of the form $K = \{h \ge \alpha\}$, with $h \in \mathcal{H}^+$ and a constant $\alpha > 0$. Notice that $K = \{h' = 1\}$ and $K \le h' \le 1$, where $h' = 1 \wedge (h/\alpha)$. Let $\mathcal{K}$ denote the $\cap c$-closure of $\mathcal{K}_0$. That is, a set K in $\mathcal{K}$ has a representation

<9> $$K = \bigcap\nolimits_{i\in\mathbb{N}} \{h_i \ge \alpha_i\}.$$

The sets in $\mathcal{K}$ are precisely those whose indicator functions are limits of decreasing sequences of functions in $\mathcal{H}^+$. The class $\mathcal{K}$ plays a role similar to that of the compact sets for measures on $\mathcal{B}(\mathbb{R}^k)$. In particular, the class

$$\mathcal{F}(\mathcal{K}) = \{F \subseteq \mathcal{X} : F \cap K \in \mathcal{K} \text{ for all } K \in \mathcal{K}\}$$

has properties analogous to the closed sets, and

$$\mathcal{B}(\mathcal{K}) = \text{sigma-field generated by } \mathcal{F}(\mathcal{K})$$

is analogous to the Borel sigma-field. Each member of $\mathcal{H}^+$ is $\mathcal{B}(\mathcal{K})$-measurable.

<10> **Theorem.** *Let $\mathcal{H}^+$ be a lattice cone, and $T : \mathcal{H}^+ \to \mathbb{R}^+$ be a map for which*

(T1) for nonnegative real numbers α_1, α_2 and functions h_1, h_2 in $\mathcal{H}^+$, $T(\alpha_1 h_1 + \alpha_2 h_2) = \alpha_1 T h_1 + \alpha_2 T h_2$;

(T2) if $h_1 \le h_2$ pointwise then $Th_1 \le Th_2$.

(T3) $Th_n \downarrow 0$ whenever the sequence $\{h_n\}$ in $\mathcal{H}^+$ decreases pointwise to zero.

(T4) $T(h \wedge n) \to Th$ as $n \to \infty$, for each h in $\mathcal{H}^+$.

Then the set function defined by

$$\mu K := \inf\{Th : K \le h \in \mathcal{H}^+\} \qquad \text{for } K \in \mathcal{K},$$
$$\mu B := \sup\{\mu K : B \supseteq K \in \mathcal{K}\}$$

is a countably additive measure on $\mathcal{B}(\mathcal{K})$ for which $Th = \mu h$ for all h in $\mathcal{H}^+$.

4. Problems

[1] Let $\{X_i : 0 \le i \le n\}$ be a submartingale with $X_0 \equiv 0$. For a fixed $\lambda \in \mathbb{R}^+$, define stopping times $\sigma := \min\{i : X_i \le -\lambda\} \wedge 1$ and $\tau := \min\{i : X_i \ge \lambda\} \wedge 1$.

(i) Show that

$$\lambda \mathbb{P}\{\max_i X_i > \lambda\} \le \mathbb{P} X_\tau \{X_\tau \ge \lambda\} \le \mathbb{P} X_1 \{X_\tau \ge \lambda\} \le \mathbb{P}|X_n|.$$

(ii) Show that

$$\lambda \mathbb{P}\{\min_i X_i < -\lambda\} \le \mathbb{P}(-X_\sigma)\{X_\sigma \le -\lambda\}$$
$$\le -\mathbb{P} X_\sigma + \mathbb{P} X_n \{X_\sigma > -\lambda\} \le \mathbb{P}|X_n|.$$

(iii) Suppose $\{Y_t : 0 \le t \le 1\}$ is a cadlag submartingale with $Y_0 \equiv 0$. Show that $\lambda \mathbb{P}\{\sup_t |Y_t| > \lambda\} \le 2\mathbb{P}|Y_1|$.

[2] Suppose a cadlag martingale $\{S_t : 0 \le t \le 1\}$, with $S_0 \equiv 0$, has a Doléans measure μ in the sense of Definition <2>, that is, $\mu((0, \tau]] = \mathbb{P} S_\tau$ for every $\tau \in \mathcal{T}_1$. Show that S has property [D] by following these steps.

(i) For a given $\tau \in \mathcal{T}_1$, let τ_n be the stopping time obtained by rounding up to the next integer multiple of 2^{-n}.

(ii) Invoke the Stopping Time Lemma to show that $0 \le \mathbb{P} S_{\tau_n}$ and $\mathbb{P} S_{\tau_n}^+ \le \mathbb{P} S_1^+$ for each $\tau \in \mathcal{T}_1$. Deduce that $\mathbb{P}|S_{\tau_n}| \le \kappa := 2\mathbb{P} S_1^+ < \infty$.

(iii) Invoke Fatou's lemma to show that $\sup_{\tau \in \mathcal{T}_1} \mathbb{P}|S_\tau| \le \kappa$.

(iv) For each $C \in \mathbb{R}^+$, show that

$$\mathbb{P} S_{\tau_n}\{S_{\tau_n} > C\} \le \mathbb{P} S_1 \{S_{\tau_n} > C\} \le \mathbb{P} S_1 \{S_1 > \sqrt{C}\} + \kappa/\sqrt{C}.$$

Invoke Fatou, then deduce that $\sup_{\tau \in \mathcal{T}_1} \mathbb{P} S_\tau \{S_\tau > C\} \to 0$ as $C \to \infty$.

(v) Show that every cadlag function on $[0, 1]$ is bounded in absolute value. Deduce that the stopping time $\sigma_C := \inf\{t : S_t < -C\} \wedge 1$ has $\sigma_C(\omega) = 1$ for all C large enough (depending on ω). Deduce that $\mu((\sigma_C, 1]] \to 0$ as $C \to \infty$.

(vi) For a given $\tau \in \mathcal{T}_1$ and $C \in \mathbb{R}^+$, define $F_\tau := \{S_\tau < -C\}$. Show that $\tau' := \tau F_\tau^c + F_\tau$ is a stopping time for which

$$\mathbb{P}\left(S_1 - S_\tau\right) F_\tau = \mathbb{P}\left(S_{\tau'} - S_\tau\right) = \mu((\tau, \tau']] \leq \mu((\sigma_C, 1]],$$

Hint: Show that if $\omega \in F_\tau$ then $\sigma_C(\omega) \leq \tau(\omega)$ and if $\omega \in F_\tau^c$ then $\tau(\omega) = \tau'(\omega)$.

(vii) Deduce that $\sup_{\tau \in \mathcal{T}_1} \mathbb{P}\left(-S_\tau\right)\{S_\tau < -C\} \to 0$ as $C \to \infty$.

[3] Let $\{S_t : t \in \mathbb{R}^+\}$ be a submartingale of class [D]. Show that there exists an integrable random variable S_∞ for which $\mathbb{P}\left(S_\infty \mid \mathcal{F}_t\right) \geq S_t \to S_\infty$ almost surely and in L^1 by following these steps.

(i) Show that the uniformly integrable submartingale $\{S_n : n \in \mathbb{N}\}$ converges almost surely and in L^1 to an S_∞ for which $\mathbb{P}\left(S_\infty \mid \mathcal{F}_n\right) \geq S_n$.

(ii) For $t \leq n$, show that $S_t \leq \mathbb{P}\left(\mathbb{P}(S_\infty \mid \mathcal{F}_n) \mid \mathcal{F}_t\right) = \mathbb{P}\left(S_\infty \mid \mathcal{F}_t\right)$.

(iii) For $t \geq n$, show that

$$\mathbb{P}(S_t - S_n)^- \leq \mathbb{P}(S_t - S_n)^+ \leq \mathbb{P}(S_\infty - S_n)^+ \to 0 \qquad \text{as } n \to \infty.$$

(iv) For each $k \in \mathbb{N}$, choose $n(k)$ for which $\mathbb{P}|S_\infty - S_{n(k)}| \leq 4^{-k}$. Invoke Problem [1] to show that $\sum_k \mathbb{P}\{\sup_{t \geq n(k)} |S_t - S_{n(k)}| > 2^{-k}\} < \infty$.

(v) Deduce that $S_t \to S_\infty$ almost surely.

5. Notes

My exposition in this Chapter is based on ideas drawn from a study of Métivier (1982, §13), Dellacherie & Meyer (1982, Chapter VII), and Chung & Williams (1990, Chapter 2). The construction in Section 2 appears new, although it is clearly closely related to existing methods.

References

Chung, K. L. & Williams, R. J. (1990), *Introduction to Stochastic Integration*, Birkhäuser, Boston.

Dellacherie, C. & Meyer, P. A. (1982), *Probabilities and Potential B: Theory of Martingales*, North-Holland, Amsterdam.

Métivier, M. (1982), *Semimartingales: A Course on Stochastic Processes*, De Gruyter, Berlin.

Pollard, D. (2001), *A User's Guide to Measure Theoretic Probability*, Cambridge University Press.

ANALYTIC SETS

For a discrete-time process $\{X_n\}$ adapted to a filtration $\{\mathcal{F}_n : n \in \mathbb{N}\}$, the prime example of a stopping time is $\tau = \inf\{n \in \mathbb{N} : X_n \in B\}$, the first time the process enters some Borel set B. For a continuous-time process $\{X_t\}$ adapted to a filtration $\{\mathcal{F}_t : t \in \mathbb{R}^+\}$, it is less obvious whether the analogously defined random variable $\tau = \inf\{t : X_t \in B\}$ is a stopping time. (Also it is not necessarily true that X_τ is a point of B.) The most satisfactory resolution of the underlying measure-theoretic problem requires some theory about analytic sets. What follows is adapted from Dellacherie & Meyer (1978, Chapter III, paras 1–33, 44–45). The following key result will be proved in this handout.

<1> **Theorem.** *Let A be a $\mathcal{B}(\mathbb{R}^+) \otimes \mathcal{F}$-measurable subset of $\mathbb{R}^+ \times \Omega$ and let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space. Then:*

(i) The projection $\pi_\Omega A := \{\omega \in \Omega : (t, \omega) \in A$ for some t in $\mathbb{R}^+\}$ belongs to $\mathcal{F}$.

(ii) There exists an $\mathcal{F}$-measurable random variable $\psi : \Omega \to \mathbb{R}^+ \cup \{\infty\}$ such that $\psi(\omega) < \infty$ and $(\psi(\omega), \omega) \in A$ for almost all ω in the projection $\pi_\Omega A$, and $\psi(\omega) = \infty$ for $\omega \notin \pi_\Omega A$.

REMARK. The map ψ in (ii) is called a ***measurable cross-section*** of the set A. Note that the cross-section $A_\omega := \{t \in \mathbb{R}^+ : (t, \omega) \in A\}$ is empty when $\omega \notin \pi_\Omega A$. It would be impossible to have $(\psi(\omega), \omega) \in A$ for such an ω.

The proofs will exploit the properties of the collection of analytic subsets of $[0, \infty] \times \Omega$. As you will see, the analytic sets have properties analogous to those of sigma-fields—stability under the formation of countable unions and intersections. They are not necessarily stable under complements, but they do have an extra stability property for projections that is not shared by measurable sets. The Theorem is made possible by the fact that the product-measurable subsets of $\mathbb{R}^+ \times \Omega$ are all analytic.

1. Notation

A collection $\mathcal{D}$ of subsets of a set $\mathcal{X}$ with $\emptyset \in \mathcal{D}$ is called a ***paving*** on $\mathcal{X}$. A paving that is closed under the formation of unions of countable subcollections is said to be a $\cup c$-paving. For example, the set $\mathcal{D}_\sigma$ of all unions of countable subcollections of $\mathcal{D}$ is a $\cup c$-paving. Similarly, the set $\mathcal{D}_\delta$ of all intersections of countable subcollections of $\mathcal{D}$ is a $\cap c$-paving. Note that $\mathcal{D}_{\sigma\delta} := (\mathcal{D}_\sigma)_\delta$ is a $\cap c$-paving but it need not be stable under $\cup c$.

Let T be a compact metric space equipped with the paving $\mathcal{K}(T)$ of compact subsets and its Borel sigma-field $\mathcal{B}(T)$, which is generated by $\mathcal{K}(T)$.

REMARK. In fact, $\mathcal{K}(T)$ is also the class of closed subsets of the compact T.

For Theorem <1>, the appropriate space will be $T = [0, \infty]$. The sets in $\mathcal{B}(\mathbb{R}^+) \otimes \mathcal{F}$ can be identified with sets in $\mathcal{B}(T) \otimes \mathcal{F}$. The compactness of T will be needed to derive good properties for the projection map $\pi_\Omega : T \times \Omega \to \Omega$.

An important role will be played by the $\cap f$-paving

$$\mathcal{K}(T) \times \mathcal{F} := \{K \times F : K \in \mathcal{K}(T),\ F \in \mathcal{F}\} \qquad \text{on } T \times \Omega$$

and by the paving $\mathcal{R}$ that consists of all finite unions of sets from $\mathcal{K}(T) \times \mathcal{F}$. That is, $\mathcal{R}$ is the $\cup f$-closure of $\mathcal{K}(T) \times \mathcal{F}$. Note (Problem [1]) that $\mathcal{R}$ is a

$(\cup f, \cap f)$-paving on $T \times \Omega$. Also, if $R = \cup_i K_i \times F_i$ then, assuming we have discarded any terms for which $K_i = \emptyset$,

$$\pi_\Omega(R) = \cup_i \pi_\Omega \left(K_i \times F_i\right) = \cup_i F_i \in \mathcal{F}.$$

REMARK. If $\mathcal{E}$ and $\mathcal{F}$ are sigma-fields, note the distinction between

$$\mathcal{E} \times \mathcal{F} = \{E \times F : E \in \mathcal{E},\ F \in \mathcal{F}\}$$

and $\mathcal{E} \otimes \mathcal{F} := \sigma(\mathcal{E} \times \mathcal{F})$.

2. Why compact sets are needed

Many of the measurability difficulties regarding projections stem from the fact that they do not "preserve set-theoretic operations" in the way that inverse images do: $\pi_\Omega\left(\cup_i A_i\right) = \cup_i \pi_\Omega A_i$ but $\pi_\Omega\left(\cap_i A_i\right) \subseteq \cap_i \pi_\Omega A_i$. Compactness of cross-sections will allow us to strengthen the last inclusion to an equality.

<2> **Lemma.** *[Finite intersection property] Suppose $\mathcal{K}_0$ is a collection of compact subsets of a metric space $\mathcal{X}$ for which each finite subcollection has a nonempty intersection. Then $\cap \mathcal{K}_0 \neq \emptyset$.*

also works for any Hausdorff topological space

Proof. Arbitrarily choose a K_0 from $\mathcal{K}_0$. If $\cap\mathcal{K}_0$ were empty then the sets $\{K^c : K \in \mathcal{K}_0\}$ would be an open cover of K_0. Extract a finite subcover $\cup_{i=1}^m K_i^c$. Then $\cap_{i=0}^m K_i = \emptyset$, a contradiction. □

<3> **Corollary.** *Suppose $\{A_i : i \in \mathbb{N}\}$ is a decreasing sequence of subsets of $T \times \Omega$ for which each ω-cross-section $K_i(\omega) := \{t \in T : (t, \omega) \in A_i\}$ is compact. Then $\pi_\Omega\left(\cap_{i\in\mathbb{N}} A_i\right) = \cap_{i\in\mathbb{N}} \pi_\Omega A_i$.*

Proof. Suppose $\omega \in \cap_{i\in\mathbb{N}} \pi_\Omega A_i$. Then $\{K_i(\omega) : i \in \mathbb{N}\}$ is a decreasing sequence of compact, nonempty (because $\omega \in \pi_\Omega A_i$) subsets of T. The finite intersection property of compact sets ensures that there is a t in $\cap_{i\in\mathbb{N}} K_i(\omega)$. The point (t, ω) belongs to $\cap_{i\in\mathbb{N}} A_i$ and $\omega \in \pi_\Omega\left(\cap_i A_i\right)$. □

REMARK. For our applications, we will be dealing only with sequences, but the argument also works for more general collections of sets with compact cross-sections.

<4> **Corollary.** *If $B = \cap_{i\in\mathbb{N}} R_i$ with $R_i \in \mathcal{R}$ then $\pi_\Omega B = \cap_{i\in\mathbb{N}} \pi_\Omega R_i \in \mathcal{F}$.*

Proof. Note that the cross-section of each $\mathcal{R}$-set is a finite union of compact sets, which is compact. Without loss of generality, we may assume that $R_1 \supseteq R_2 \supseteq \ldots$. Invoke Corollary <3>. □

3. Measurability of some projections

For which $B \in \mathcal{B}(T) \otimes \mathcal{F}$ is it true that $\pi_\Omega(B) \in \mathcal{F}$? From Corollary <4>, we know that it is true if B belongs to $\mathcal{R}_\delta$. The following properties of outer measures (see Problem [2]) will allow us to extend this nice behavior to sets in $\mathcal{R}_{\sigma\delta}$:

(i) If $A_1 \subseteq A_2$ then $\mathbb{P}^*(A_1) \leq \mathbb{P}^*(A_2)$

(ii) If $\{A_i : i \in \mathbb{N}\}$ is an increasing sequence then $\mathbb{P}^*\left(A_i\right) \uparrow \mathbb{P}^*\left(\cup_{i\in\mathbb{N}} A_i\right)$.

(iii) If $\{F_i : i \in \mathbb{N}\} \subseteq \mathcal{F}$ is a decreasing sequence then

$$\mathbb{P}^*\left(F_i\right) = \mathbb{P}F_i \downarrow \mathbb{P}\left(\cap_{i\in\mathbb{N}} F_i\right) = \mathbb{P}^*\left(\cap_{i\in\mathbb{N}} F_i\right).$$

For each subset D of $T \times \Omega$ define $\Psi^*(D) := \mathbb{P}^* \pi_\Omega D$, the outer measure of the projection of D onto Ω. If $D_i \uparrow D$ then $\pi_\Omega D_i \uparrow \pi_\Omega D$. If $R_i \in \mathcal{R}$ and

$R_i \downarrow B$ then $\pi_\Omega R_i \in \mathcal{F}$ and $\pi_\Omega R_i \downarrow \pi_\Omega B \in \mathcal{F}$. The properties for $\mathbb{P}^*$ carry over to analogous properties for Ψ^*:

(i) If $D_1 \subseteq D_2$ then $\Psi^*(D_1) \le \Psi^*(D_2)$

(ii) If $\{D_i : i \in \mathbb{N}\}$ is an increasing sequence then $\Psi^*\left(D_i\right) \uparrow \Psi^*\left(\cup_{i\in\mathbb{N}} D_i\right)$.

(iii) If $\{R_i : i \in \mathbb{N}\} \subseteq \mathcal{R}$ is a decreasing sequence then $\Psi^*\left(R_i\right) \downarrow \Psi^*\left(\cap_{i\in\mathbb{N}} R_i\right)$.

With just these properties, we can show that π_Ω behaves well on a much larger collection of sets than $\mathcal{R}$.

<5> **Lemma.** *If $A \in \mathcal{R}_{\sigma\delta}$ then $\Psi^*(B) = \sup\{\Psi^*(B) : B \in \mathcal{R}_\delta\}$. Consequently, the set $\pi_\Omega A$ belongs to $\mathcal{F}$.*

Proof. Write A as $\cap_{i\in\mathbb{N}} D_i$ with $D_i = \cup_{j\in\mathbb{N}} R_{ij} \in \mathcal{R}_\sigma$. As $\mathcal{R}$ is $\cup f$-stable, we may assume that R_{ij} is increasing in j for each fixed i.

Suppose $\Psi^*(A) > M$ for some constant M. Invoke (ii) for the sequence $\{AR_{1j}\}$, which increases to $AD_1 = A$, to find an index j_1 for which the set $R_1 := R_{1j_1}$ has $\Psi^*\left(AR_1\right) > M$.

The sequence $\{AR_1R_{2j}\}$ increases to $AR_1D_2 = AR_1$. Again by (ii), there exists an index j_2 for which the set $R_2 = R_{2j_2}$ has $\Psi^*(AR_1R_2) > M$. And so on. In this way we construct sets R_i in $\mathcal{R}$ for which

$$\Psi^*(R_1R_2 \ldots R_n) \ge \Psi^*(AR_1R_2 \ldots R_n) > M$$

for every n. The set $B_M := \cap_{i\in\mathbb{N}} R_i$ belongs to $\mathcal{R}_\delta$; it is a subset of $\cap_{i\in\mathbb{N}} D_i = A$; and, by (iii), $\Psi^*(B) \ge M$.

By Corollary <4>, the set B_M projects to a set $F_M := \pi_\Omega B_M$ in $\mathcal{F}$ and hence $\mathbb{P}F_M = \Psi^* B \ge M$. The set $\pi_\Omega A$ is inner regular, in the sense that

$$\mathbb{P}^*\pi_\Omega A = \Psi^* A = \sup\{\mathbb{P}F : \pi_\Omega A \supseteq F \in \mathcal{F}\}$$

□ It follows (Problem [2]) that the set $\pi_\Omega A$ belongs to $\mathcal{F}$.

The properties shared by $\mathbb{P}^*$ and Ψ^* are so useful that they are given a name.

<6> **Definition.** *Suppose $\mathcal{S}$ is a paving on a set S. A function Ψ defined for all subsets of S and taking values in $[-\infty, \infty]$ is said to be a **Choquet $\mathcal{S}$-capacity** if it satisfies the following three properties.*

(i) If $D_1 \subseteq D_2$ then $\Psi(D_1) \le \Psi(D_2)$

(ii) If $\{D_i : i \in \mathbb{N}\}$ is an increasing sequence then $\Psi\left(D_i\right) \uparrow \Psi\left(\cup_{i\in\mathbb{N}} D_i\right)$.

(iii) If $\{S_i : i \in \mathbb{N}\} \subseteq \mathcal{S}$ is a decreasing sequence then $\Psi\left(S_i\right) \downarrow \Psi\left(\cap_{i\in\mathbb{N}} S_i\right)$.

The outer measure $\mathbb{P}^*$ is a Choquet $\mathcal{F}$-capacity defined for the subsets of Ω. Moreover, if Ψ is any Choquet $\mathcal{F}$-capacity defined for the subsets of Ω then $\Psi^*(D) := \Psi(\pi_\Omega D)$ is a Choquet $\mathcal{R}$-capacity defined for the subsets of $T \times \Omega$. The argument from Lemma <5> essentially shows that if $A \in \mathcal{R}_{\sigma\delta}$ then $\Psi^*(B) = \sup\{\Psi^*(B) : B \in \mathcal{R}_\delta\}$ for every such Ψ^*, whether defined via $\mathbb{P}^*$ or not.

4. Analytic sets

The paving of $\mathcal{S}$-analytic sets can be defined for any paving $\mathcal{S}$ on a set S. For our purposes, the most important case will be $S = T \times \Omega$ with $\mathcal{S} = \mathcal{R}$.

<7> **Definition.** *Suppose $\mathcal{S}$ is a paving on a set S. A subset A of S is said to be $\mathcal{S}$-analytic if there exists a compact metric space E and a subset D in $\left(\mathcal{K}(E) \times \mathcal{S}\right)_{\sigma\delta}$ for which $A = \pi_S D$. Write $\mathcal{A}(\mathcal{S})$ for the set of all $\mathcal{S}$-analytic subsets of S.*

REMARK. Note that $\mathcal{R}_{\sigma\delta} = (\mathcal{K}(T) \times \mathcal{F})_{\sigma\delta}$. The σ takes care of the $\cup f$ operation needed to generate $\mathcal{R}$ from $\mathcal{K}(T) \times \mathcal{F}$. The $\mathcal{R}$-analytic sets are also called $\mathcal{K}(T) \times \mathcal{F}$-analytic sets.

In fact, it is possible to find a single E that defines all the $\mathcal{S}$-analytic subsets, but that possibility is not important for my purposes. What is important is the fact that $\mathcal{A}(\mathcal{S})$ is a $(\cup c, \cap c)$-paving: see Problem [3].

When E is another compact metric space, Tychonoff's theorem (see Dudley 1989, Section 2.2, for example) ensures not only that the product space $E \times T$ is a compact metric space but also that $\mathcal{K}(E) \times \mathcal{K}(T) \subseteq \mathcal{K}(E \times T)$.

Lemma <5>, applied to $\widetilde{T} := E \times T$ instead of T and with $\widetilde{\mathcal{R}}$ the $\cup f$-closure of $\mathcal{K}(E \times T) \times \mathcal{F}$, implies that

<8> $$\widetilde{\pi}_\Omega D \in \mathcal{F} \qquad \text{for each } D \text{ in } \widetilde{\mathcal{R}}_{\sigma\delta}.$$

Here $\widetilde{\pi}_\Omega$ projects $E \times T \times \Omega$ onto Ω. We also have

$$\widetilde{\mathcal{R}}_{\sigma\delta} \supseteq \left(\mathcal{K}(E) \times \mathcal{K}(T) \times \mathcal{F}\right)_{\sigma\delta} = \left(\mathcal{K}(E) \times \mathcal{R}\right)_{\sigma\delta}$$

where $\mathcal{R}$ is the $\cup f$-closure of $\mathcal{K}(T) \times \mathcal{F}$, as in Section 3. As a special case of property <8> we have

<9> $$\widetilde{\pi}_\Omega D \in \mathcal{F} \qquad \text{for each } D \text{ in } \left(\mathcal{K}(E) \times \mathcal{R}\right)_{\sigma\delta}.$$

Write $\widetilde{\pi}_\Omega$ as a composition of projection $\pi_\Omega \circ \widetilde{\pi}_{T\times\Omega}$, where $\widetilde{\pi}_{T\times\Omega}$ projects $E \times T \times \Omega$ onto $T \times \Omega$. As E ranges over all compact metric spaces and D ranges over all the $\left(\mathcal{K}(E) \times \mathcal{R}\right)_{\sigma\delta}$ sets, the projections $A := \widetilde{\pi}_{T\times\Omega} D$ range over all $\mathcal{R}$-analytic subsets of $T \times \Omega$. Property <9> is equivalent to the assertion

<10> $$\pi_\Omega A \in \mathcal{F} \qquad \text{for all } A \in \mathcal{A}(\mathcal{R}).$$

In fact, the method used to prove Lemma <5> together with an analogue of the argument just outlined establishes an approximation theorem for analytic sets and general Choquet capacities.

<11> **Theorem.** *Suppose $\mathcal{S}$ is a $(\cup f, \cap f)$-paving on a set S and Let Ψ is a Choquet $\mathcal{S}$-capacity on S. Then $\Psi(A) = \sup\{\Psi(B) : A \supseteq B \in \mathcal{S}_\delta\}$. for each A in $\mathcal{A}(\mathcal{S})$.*

To prove assertion (i) of Theorem <1>, we have only to check that

$$\mathcal{B}(T) \otimes \mathcal{F} \subseteq \mathcal{A}(\mathcal{R})$$

for the special case where $T = [0, \infty]$. By Problem [3], $\mathcal{A}(\mathcal{R})$ is a $(\cup c, \cap c)$-paving. It follows easily that

$$\mathcal{H} := \{H \in \mathcal{B}(T) \otimes \mathcal{F} : H \in \mathcal{A}(\mathcal{R}) \text{ and } H^c \in \mathcal{A}(\mathcal{R})\}$$

is a sigma-field on $T \times \Omega$. Each $K \times F$ with $K \in \mathcal{K}(T)$ and $F \in \mathcal{F}$ belongs to $\mathcal{H}$ because $\mathcal{K}(T) \times \mathcal{F} \subseteq \mathcal{R} \subseteq \mathcal{A}(\mathcal{R})$ and

$$(K \times F)^c = \left(K \times F^c\right) + \left(K^c \times \Omega\right)$$
$$K^c = \cup_{i\in\mathbb{N}}\{t : d(t, K) \geq 1/i\} \in \mathcal{K}(T)_\sigma$$

It follows that $\mathcal{H} = \sigma(\mathcal{K}(T) \times \mathcal{F}) = \mathcal{B}(T) \otimes \mathcal{F}$ and $\mathcal{B}(T) \otimes \mathcal{F} \subseteq \mathcal{A}(\mathcal{R})$.

5. Existence of measurable cross-sections

The general Theorem <11> is exactly what we need to prove part (ii) of Theorem <1>.

Once again identify A with an $\mathcal{R}$-analytic subset of $T \times \Omega$, where $T = [0, \infty]$. The result is trivial if $\alpha_1 := \mathbb{P}\pi_\Omega A = 0$, so assume $\alpha_1 > 0$.

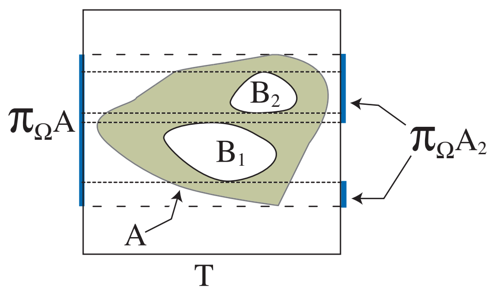
Invoke Theorem <11> for the $\mathcal{R}$-capacity defined by $\Psi^*(D) = \mathbb{P}^*(\pi_\Omega D)$. Find a subset with $A \supseteq B_1 \in \mathcal{R}_\delta$ and $\mathbb{P}\left(\pi_\Omega B_1\right) = \Psi^*(B_1) \geq \alpha_1/2$. Define

$$\psi_1(\omega) := \inf\{t \in \mathbb{R}^+ : (t, \omega) \in B_1\}.$$

Because the set B_1 has compact cross-sections, the infimum is actually achieved for each ω in $\pi_\Omega B_1$. For $\omega \notin \pi_\Omega B_1$ the infimum equals ∞. Define

$$A_2 := \{(t, \omega) \in A : \omega \notin \pi_\Omega B_1\} = A \cap \left(T \times (\pi_\Omega B_1)^c\right)$$

Note that $A_2 \in \mathcal{A}(\mathcal{R})$ and $\alpha_2 := \mathbb{P}\pi_\Omega A_2 \leq \alpha_1/2$. Without loss of generality suppose $\alpha_2 > 0$. Find a subset with $A_2 \supseteq B_2 \in \mathcal{R}_\delta$ and $\mathbb{P}\left(\pi_\Omega B_2\right) = \Psi^*(B_2) \geq \alpha_2/2$. Define $\psi_2(\omega)$ as the first hitting time on B_2.



And so on. The sets $\{\pi_\Omega B_i : i \in \mathbb{N}\}$ are disjoint, with $F := \cup_{i\in\mathbb{N}}\pi_\Omega B_i$ a subset of $\pi_\Omega A$. By construction $\alpha_i \downarrow 0$, which ensures that $\mathbb{P}\left((\pi_\Omega A)\backslash F\right) = 0$.

If $\alpha_i = 0$ for some i, the construction requires only finitely many steps.

Define $\psi := \inf_{i\in\mathbb{N}} \psi_i$. On B we have $(\psi(\omega), \omega) \in A$. □

6. Problems

[1] Suppose $\mathcal{S}$ is a paving (on a set S), which is $\cap f$-stable. Let $\mathcal{S}_{\cup f}$ consists of the set of all unions of finite collections of sets from $\mathcal{S}$. Show that $\mathcal{S}_{\cup f}$ is a $(\cup f, \cap f)$-paving. Hint: Show that $(\cup_i S_i) \cap (\cup_j T_j) = \cup_{i,j}(S_i \cap T_j)$.

[2] The outer measure of a set $A \subseteq \Omega$ is defined as $\mathbb{P}A := \inf\{\mathbb{P}F : A \subseteq F \in \mathcal{F}\}$.

(i) Show that the infimum is achieved, that is, there exists an $F \in \mathcal{F}$ for which $A \subseteq F$ and $\mathbb{P}^*A = \mathbb{P}F$. Hint: Consider the intersection of a sequence of sets for which $\mathbb{P}F_n \downarrow \mathbb{P}^*A$.

(ii) Suppose $\{D_n : n \in \mathbb{N}\}$ is an increasing sequence of sets (not necessarily $\mathcal{F}$-measurable) with union D. Show that $\mathbb{P}^*D_n \uparrow \mathbb{P}^*D$. Hint: Find sets with $D_i \subseteq F_i \in \mathcal{F}$ and $\mathbb{P}^*D_i = \mathbb{P}F_i$. Show that $\cap_{i\geq n}F_i \uparrow F \supseteq D$ and $\mathbb{P}F \leq \sup_{i\in\mathbb{N}}\mathbb{P}^*D_i$.

(iii) Suppose D is a subset of Ω for which $\mathbb{P}^*D = \sup\{\mathbb{P}F_0 : D \supseteq F_0 \in \mathcal{F}\}$. Show that D belongs to the $\mathbb{P}$-completion of $\mathcal{F}$ (or to $\mathcal{F}$ itself if $\mathcal{F}$ is $\mathbb{P}$-complete). Hint: Find sets F and F_i in $\mathcal{F}$ for which $F_i \subseteq D \subseteq F$ and $\mathbb{P}F_i \uparrow \mathbb{P}^*D = \mathbb{P}F$. Show that $F\backslash\left(\cup_{i\in\mathbb{N}} F_i\right)$ has zero $\mathbb{P}$-measure.

[3] Suppose $\{A_\alpha : \alpha \in \mathbb{N}\} \subseteq \mathcal{A}(\mathcal{S})$. Show that $\cup_\alpha A_a \in \mathcal{A}(\mathcal{S})$ and $\cap_\alpha A_a \in \mathcal{A}(\mathcal{S})$, by the following steps. Recall that there exist compact metric spaces $\{E_\alpha : \alpha \in \mathbb{N}\}$, each equipped with its paving $\mathcal{K}_\alpha$ of compact subsets, and sets $D_\alpha \in \left(\mathcal{K}_\alpha \times \mathcal{S}\right)_{\sigma\delta}$ for which $A_\alpha = \pi_S D_\alpha$.

D&M Theorem 3.8

(i) Define $E := \times_{\alpha\in\mathbb{N}} E_\alpha$ and $E_{-\beta} = \times_{\alpha\in\mathbb{N}\setminus\{\beta\}} E_\alpha$. Show that E is a compact metric space.

(ii) Define $\widetilde{D} := D_\alpha \times E_{-\alpha}$. Show that $\widetilde{D}_\alpha \in \big(\mathcal{K}(E) \times \mathcal{S}\big)_{\sigma\delta}$ and that $A_\alpha = \widetilde{\pi}_S \widetilde{D}_\alpha$, where $\widetilde{\pi}_S$ denotes the projection map from $E \times S$ to S.

(iii) Show that $\cap_\alpha A_\alpha = \widetilde{\pi}_S \left(\cap_\alpha \widetilde{D}_\alpha \right)$ and $\cap_\alpha \widetilde{D}_\alpha \in \big(\mathcal{K}(E) \times \mathcal{S}\big)_{\sigma\delta}$.

(iv) Without loss of generality suppose the E_α spaces are disjoint—otherwise replace E_α by $\{\alpha\} \times E_\alpha$. Define $H = \cup_{\alpha\in\mathbb{N}} E_\alpha$ and $E^* := H \cup \{\infty\}$. Without loss of generality suppose the metric d_α on E_α is bounded by $2^{-\alpha}$. Define

$$d(x, y) = d(y, x) := \begin{cases} d_\alpha(x, y) & \text{if } x, y \in E_\alpha \\ 2^{-\alpha} + 2^{-\beta} & \text{if } x \in E_\alpha,\ y \in E_\beta \text{ with } \alpha \neq \beta \\ 2^{-\alpha} & \text{if } y = \infty \text{ and } x \in E_\alpha \end{cases}$$

Show that E^* is a compact metric space under d.

(v) Suppose $D_\alpha = \cap_{i\in\mathbb{N}} B_{\alpha i}$ with $B_{\alpha i} \in (\mathcal{K}_\alpha \times \mathcal{S})_\sigma$. Show that $\cup_\alpha D_\alpha = \cap_i \cup_\alpha B_{\alpha,i}$. Hint: Consider the intersection with $E_\alpha \times S$.

(vi) Deduce that $\cup_\alpha D_\alpha \in (\mathcal{K}(E^*) \times \mathcal{S})_{\sigma\delta}$.

(vii) Conclude that $\cup_\alpha A_\alpha = \pi_S \cup_\alpha D_\alpha \in \mathcal{A}(\mathcal{S})$.

References

Dellacherie, C. & Meyer, P. A. (1978), *Probabilities and Potential*, North-Holland, Amsterdam. (First of three volumes).

Dudley, R. M. (1989), *Real Analysis and Probability*, Wadsworth, Belmont, Calif.

A standard Brownian motion on $\mathbb{R}^+$ for a filtration $\{\mathcal{F}_t : t \in \mathbb{R}^+\}$ is an adapted process for which

(i) all sample paths are continuous

(ii) $X(0, \omega) = 0$ for all ω

(iii) for each pair s, t with $0 \le s \le t$,

$$X_t - X_s \qquad \text{is } N(0, t-s) \text{ distributed independent of } \mathcal{F}_s$$

Properties (ii) and (iii) together imply: for each $0 \le t_1 \le t_2 \le \ldots \le t_k$, the random vector $(X_{t_1}, \ldots, X_{t_k})$ has a multivariate normal distribution with zero means and covariances given by

$$\operatorname{cov}(X_s, X_t) = \min(s, t)$$

Abbreviate $\mathbb{P}(\ldots \mid \mathcal{F}_t)$ to $\mathbb{P}_t(\ldots)$.

Useful facts: some rigorous proofs to follow

(i) For a fixed $\tau \ge 0$ define $Z_t = B_{\tau+t} - B_\tau$ for $t \ge 0$. Then Z is a standard Brownian motion independent of $\mathcal{F}_\tau$.

(ii) (Strong Markov property) Same assertion as in (i) except that τ is a stopping time.

(iii) (Time reversal) Define $Z_t = tB_{1/t}$ for $t > 0$, with $Z_0 = 0$. Then $\{Z_t : t \in \mathbb{R}^+\}$ is a also a standard Brownian motion.

(iv) Both $\{(B_t, \mathcal{F}_t) : t \in \mathbb{R}^+\}$ and $\{(B_t^2 - t, \mathcal{F}_t) : t \in \mathbb{R}^+\}$ are martingales.

(v) For each real θ, the process $Y_t = \exp\left(\theta X_t - \frac{1}{2}\theta^2 t\right)$ is a martingale. (For complex θ, would it be a complex martingale?)

<1> **Lévy's martingale characterization of Brownian motion.** *Suppose $\{X_t : 0 \le t \le 1\}$ is a martingale with continuous sample paths and $X_0 = 0$. Suppose also that $X_t^2 - t$ is a martingale. Then X is a Brownian motion.*

Heuristics of the proof that $X_1 \sim N(0, 1)$. The two martingale assuptions give two properties of the increment $\Delta X = X_t - X_s$, for $s < t$:

<2> $$\mathbb{P}_s \Delta X = 0 \qquad \text{and} \qquad \mathbb{P}_s (\Delta X)^2 = t - s.$$

Let $f(x, t)$ be a smooth function of two arguments, $x \in \mathbb{R}$ and $t \in [0, 1]$. Define

$$f_x = \frac{\partial f}{\partial x} \qquad \text{and} \qquad f_{xx} = \frac{\partial^2 f}{\partial^2 x} \qquad \text{and} \qquad f_t = \frac{\partial f}{\partial t}.$$

Let $h = 1/n$ for some large positive integer n. Define $t_i = ih$ for $i = 0, 1, \ldots, n$. Write $\Delta_i X$ for $X(t_i + h) - X(t_i)$. Then

$$\begin{aligned}
&\mathbb{P} f(X_1, 1) - \mathbb{P} f(X_0, 0) \\
&\qquad = \sum_{i<n} \left(\mathbb{P} f(X_{t_i+h}, t_i + h) - \mathbb{E} f(X_{t_i}, t_i) \right) \\
&\qquad \approx \sum_{i<n} \mathbb{P} \left((\Delta_i X) f_x(X_{t_i}, t_i) + \tfrac{1}{2} (\Delta_i X)^2 f_{xx}(X_{t_i}, t_i) + h f_t(X_{t_i}, t_i) \right) \\
&\qquad = \sum_{i<n} \left(0 + \tfrac{1}{2} h \mathbb{P} f_{xx}(X_{t_i}, t_i) + h \mathbb{P} f_t(X_{t_i}, t_i) \right) \\
&\qquad \approx \int_0^1 \left(\tfrac{1}{2} \mathbb{P} f_{xx}(X_s, s) + \mathbb{P} f_t(X_s, s) \right) ds \qquad \text{if } h \text{ is small} \qquad \text{by <2>.}
\end{aligned}$$

We need to formalize the passage to the limit to get

$$\mathbb{P}f(X_1, 1) - \mathbb{P}f(X_0, 0) = \int_0^1 \left(\tfrac{1}{2}\mathbb{P}f_{xx}(X_s, s) + \mathbb{P}f_t(X_s, s)\right) ds.$$

Specialize to the case $f(x, s) = \exp\left(\theta x - \frac{1}{2}\theta^2 s\right)$, with θ a fixed constant. By direct calculation,

$$f_x = \theta f(x, s) \qquad \text{and} \qquad f_{xx} = \theta^2 f(x, s) \qquad \text{and} \qquad f_t = -\tfrac{1}{2}\theta^2 f(x, s)$$

Thus

$$\mathbb{P}e^{\theta X_1} e^{-\theta^2/2} - 1 = \int_0^1 0\, ds = 0.$$

That is, X_1 has the moment generating function $\exp(\theta^2/2)$, which identifies it as having a $N(0, 1)$ distribution. □

As you will see, we are effectively proving a martingale central limit theorem. Look at the handout ***martingaleCLT.pdf*** before we start on a rigorous proof.