Introduction to Semi-martingale Theory

Josef Teichmann

Autumn term 2020

Martingales and all that

Filtered probability space

Given a probability space $(\Omega, \mathcal{F}, P)$, a *filtration* is an increasing family of sub-σ-algebras $(\mathcal{F}_t)_{t \in T}$ for a given index set $T \subset \mathbb{R} \cup \{\pm\infty\}$.

Usual conditions

We shall often assume the "usual conditions" on a filtered probability space, i.e. that a filtration is right continuous and complete, but we first name the properties separately:

1. A filtration is called *complete* if each $\mathcal{F}_t$ contains all P-null sets from $\mathcal{F}$ for $t \in T$. In particular every σ-algebra $\mathcal{F}_t$ is complete with respect to P.
2. A filtration is called *right continuous* if $\mathcal{F}_t = \mathcal{F}_{t+} := \cap_{\epsilon>0} \mathcal{F}_{t+\epsilon}$ for all $t \in T$, where $\mathcal{F}_{t+\epsilon} := \cap_{u \in T, u \geq t+\epsilon} \mathcal{F}_u$ for $\epsilon > 0$.

Stopping times

Definition

A $T \cup \{\infty\}$-valued random variable τ is called an $(\mathcal{F}_t)_{t\in T}$-stopping time or $(\mathcal{F}_t)_{t\in T}$-optional time if $\{\tau \leq t\} \in \mathcal{F}_t$ for all $t \in T$.
A $T \cup \{\infty\}$-valued random variable τ is called a weak $(\mathcal{F}_t)_{t\in T}$-stopping time or a weakly $(\mathcal{F}_t)_{t\in T}$-optional time if $\{\tau < t\} \in \mathcal{F}_t$ for all $t \in T$.

We can collect several results:

1. Let $(\mathcal{F}_t)_{t\in T}$ be a right continuous filtration, then every weakly optional time is optional.
2. Suprema of sequences of optional times are optional, infima of sequences of weakly optional times are weakly optional.
3. Any weakly optional time τ taking values in $\mathbb{R} \cup \{\pm\infty\}$ can be approximated by some countably valued optional time $\tau_n \searrow \tau$, take for instance $\tau_n := 2^{-n}[2^n\tau + 1]$.

Stochastic Processes

A stochastic process is a family of random variables $(X_t)_{t \in T}$ on a filtered probability space $(\Omega, \mathcal{F}, P)$. The process is said to be $(\mathcal{F}_t)_{t \in T}$-adapted if X_t is $\mathcal{F}_t$-measurable for all $t \in T$.

Two stochastic processes $(X_t)_{t \in T}$ and $(Y_t)_{t \in T}$ are called *versions* of each other if $X_t = Y_t$ almost surely for all $t \in T$, i.e. for every $t \in T$ there is a null set N_t such that $X_t(\omega) = Y_t(\omega)$ for $\omega \notin N_t$.

Two stochastic processes $(X_t)_{t \geq 0}$ and $(Y_t)_{t \geq 0}$ are called *indistinguishable* if almost everywhere $X_t = Y_t$ for all $t \in T$, i.e. there is a set N such that for all $t \in T$ equality $X_t(\omega) = Y_t(\omega)$ holds true for $\omega \notin N$.

Path properties

A stochastic process is called cadlag or RCLL (caglad or LCRL) if the sample paths $t \mapsto X_t(\omega)$ are right continuous and have limits at the left side (left continuous and have limits at the right hand side) for all $\omega \in \Omega$. Given a stochastic process $(X_t)_{t\in T}$ we denote by $\hat{X}$ the associated mapping from $T \times \Omega$ to $\mathbb{R}$, which maps (t, ω) to $X_t(\omega)$.

Progressively measurable

Given a general filtration: a stochastic process $(X_t)_{t \in T}$ is called progressively measureable or progressive if $\hat{X} : ((T \cap [-\infty, t]) \times \Omega, \mathcal{B}(T \cap [-\infty, t]) \otimes \mathcal{F}_t) \to \mathbb{R}$ is measureable for all t.

An adapted stochastic process with right continuous paths (or left continuous paths) almost surely is progressively measureable.

A stochastic process $(X_t)_{t \geq 0}$ is called measureable if $\hat{X}$ is measureable with respect to the product σ-algebra. A measureable adapted process has a progressively measureable modification, which is a complicated result.

Given a general filtration we can consider an associated filtration on the convex hull $\overline{T}$ of T, where right from right-discrete points (i.e. $t \in T$ such that there is $\delta > 0$ with $]t, t+\delta[\cap T = \emptyset$), a right continuous, piece-wise constant extension is taken.

We introduce two further σ-algebras on $\overline{T} \times \Omega$, namely the one generated by left continuous processes (the predictable σ-algebra $\mathcal{P}$) and the one generated by the right continuous processes (the optional σ-algebra $\mathcal{O}$). Obviously $\mathcal{P} \subset \mathcal{O}$.

Given a stochastic processes $(X_t)_{t \in T}$ then we call it predictable if there is an extension on $\overline{T}$ measurable with respect to $\mathcal{P}$.

Most important stopping times are hitting times of stochastic processes $(X_t)_{t\in T}$: given a Borel set B we can define the hitting time τ of a stochastic process $(X_t)_{t\in T}$ via

$$\tau = \inf\{t \in T \mid X_t \in B\}\,.$$

There are several profound results around hitting times useful in potential theory, we provide some simple ones:

1. Let $(X_t)_{t \in T}$ be an $(\mathcal{F}_t)_{t \in T}$-adapted right continuous process with respect to a general filtration and let B be an open set, then the hitting time is a weak $(\mathcal{F}_t)_{t \in T}$-stopping time.
2. Let $(X_t)_{t \in T}$ be an $(\mathcal{F}_t)_{t \in T}$-adapted continuous process with respect to a general filtration and let B be a closed set, then the hitting time is a $(\mathcal{F}_t)_{t \in T}$-stopping time.
3. It is a very deep result in potential theory that all hitting times of Borel subsets are $(\mathcal{F}_t)_{t \in T}$-stopping times, if usual conditions are satisfied and the process is progressively measurable (Debut theorem).

The σ-algebra $\mathcal{F}_t$ represents the set of all (theoretically) observable events up to time t including t, the stopping time σ-algebra $\mathcal{F}_\tau$ represents all (theoretically) observable events up to τ:

$$\mathcal{F}_\tau = \{A \in \mathcal{F}, A \cap \{\tau \leq t\} \in \mathcal{F}_t \text{ for all } t \in T\}\,.$$

Martingales

Definition

A stochastic process $(X_t)_{t \in T}$ is called martingale (submartingale, supermartingale) with respect to a filtration $(\mathcal{F}_t)_{t \in T}$ if

$$E[X_t \mid \mathcal{F}_s] = X_s$$

for $t \geq s$ in T ($E[X_t \mid \mathcal{F}_s] \geq X_s$ for submartingales, $E[X_t \mid \mathcal{F}_s] \leq X_s$ for super-martingales).

Proposition

Let M be a martingale on an arbitrary index set T with respect to a filtration $(\mathcal{G}_t)_{t \in T}$. Assume a second filtration $(\mathcal{F}_t)_{t \in T}$ such that $\mathcal{F}_t \subset \mathcal{G}_t$ for $t \in T$ and assume that M is actually also $\mathcal{F}$ adapted, then M is also a martingale with respect to the filtration $(\mathcal{F}_t)_{t \in T}$.

Doob-Meyer decomposition in discrete time

Remember the Doob-Meyer decomposition for any integrable adapted stochastic process $(X_t)_{t\in\mathbb{N}}$ on the *countable index set* $T=\mathbb{N}$: there is a unique martingale M and a unique predictable process A with $A_0=0$ such that $X=M-A$. It can be defined directly via

$$A_t := \sum_{s<t} E[X_s - X_{s+1} \mid \mathcal{F}_s]$$

for $t\in\mathbb{N}$, and $X=M-A$. The decomposition is unique since a predictable martingale starting at 0 vanishes. If X is a super-martingale, then the Doob-Meyer decomposition $X=M-A$ yields a martingale and an increasing process A.

Estimates of Doob-Meyer decomposition

We shall need in the sequel a curious inequality for bounded, non-negative super-martingales. We state first a lemma for super-martingales of the type $X_s = E[A_\infty|\mathcal{F}_s] - A_s$, for $s \geq 0$, where A is a non-negative, predictable, increasing process with $A_0 = 0$ and limit at infinity A_∞ (such a super-martingale is called a potential). Assume $0 \leq X \leq c$, then

$$E[A_\infty^p] \leq p! c^{p-1} E[X_0] \,,$$

for natural numbers $p \geq 1$.

Proof

Indeed

$$
\begin{aligned}
A_\infty^p &= \sum_{i_1,\dots,i_p} (A_{i_1+1} - A_{i_1}) \dots (A_{i_p+1} - A_{i_p}) \\
&= p \sum_j \sum_{j_1,\dots,j_{p-1} \le j} (A_{j_1+1} - A_{j_1}) \dots (A_{j_{p-1}+1} - A_{j_{p-1}})(A_{j+1} - A_j) \\
&= p \sum_{j_1,\dots,j_{p-1}} (A_{j_1+1} - A_{j_1}) \dots (A_{j_{p-1}+1} - A_{j_{p-1}})(A_\infty - A_{j_1 \vee \dots \vee j_{p-1}})
\end{aligned}
$$

by changing the order of summation.

Proof

Next we take the conditional expectations on the information $\mathcal{F}_{j_1 \vee \ldots \vee j_{p-1}}$, observing that all terms (predictability!) except the last one are measurable with respect to $\mathcal{F}_{j_1 \vee \ldots \vee j_{p-1}}$ and inserting

$$X_{j_1 \vee \ldots \vee j_{p-1}} = E[A_\infty - A_{j_1 \vee \ldots \vee j_{p-1}} \mid \mathcal{F}_{j_1 \vee \ldots \vee j_{p-1}}] \leq c$$

we obtain the recursive inequality

$$E[A_\infty^p] \leq pE[A_\infty^{p-1}]c\,,$$

which leads by induction to the desired result, since $E[A_\infty] = E[X_0]$ by the definition of a potential.

Let us now consider a non-negative super-martingale X bounded by a constant $c \geq 0$, then we can apply the previous inequality for the potential $X_0, \ldots, X_n, X_\infty, 0, 0, \ldots$ (notice here that also the filtration changes)with corresponding $\tilde{A}^{(n)}_\infty = A_n + E[X_n - X_\infty \mid \mathcal{F}_n] + X_\infty \geq A_n + X_\infty$. If we let n tend to ∞ (notice the almost sure convergence of the bounded super-martingale, see below), we obtain the result, namely that

$$E[M^p_\infty] \leq p! c^{p-1} E[X_0] \,, \tag{1}$$

since $M_\infty = A_\infty + X_\infty = \lim_{n \to \infty} \tilde{A}^{(n)}_\infty$.

Optional Sampling: the discrete case

The most important theorem of this section is Doob's optional sampling theorem: it states that the stochastic integral, often also called martingale transform, with respect to a martingale is again a martingale.

Let us consider a finite index set T, a stochastic process $X = (X_t)_{t \in T}$ and an increasing sequence of stopping times $(\tau_k)_{k \geq 0}$ taking values in T together with bounded random variables V_k, which are $\mathcal{F}_{\tau_k}$, then

$$V_t := \sum_{k \geq 0} V_k 1_{\{\tau_k < t \leq \tau_{k+1}\}}$$

is a predictable process. We call such processes *simple predictable* and we can define the stochastic integral (as a finite sum)

$$(V \bullet X)_{s,t} := \sum_{k \geq 0} V_k (X_{s \vee t \wedge \tau_{k+1}} - X_{s \vee t \wedge \tau_k}) \,.$$

By simple conditioning arguments we can prove the following proposition:

Proposition

Assume that the stopping times are deterministic and let X be a martingale, then $(V \bullet X)$ is a martingale. If X is a sub-martingale and $V \geq 0$, then $(V \bullet X)$ is a sub-martingale, too. Furthermore it always holds that $(V \bullet (W \bullet X)) = (VW \bullet X)$ (Associativity).

This basic proposition can be immediately generalized by considering stopping times instead of deterministic times, since for any two stopping times $\tau_1 \leq \tau_2$ taking values in T we have

$$1_{\{\tau_1 < t \leq \tau_2\}} = \sum_{k \geq 0} 1_{\{\tau_1 < t \leq \tau_2\}} 1_{\{t_k < t \leq t_{k+1}\}}$$

for the sequence $t_0 < t_1 < t_2 < ...$ exhausting the finite set T. Whence we can argue by Associativity $(V \bullet (W \bullet X)) = (VW \bullet X)$ that we do always preserve the martingale property, or the sub-martingale property, respectively, in case of non-negative integrands.

We obtain several simple conclusions from these basic facts (optional sampling theorem):

A stochastic process X is a martingale if and only if for each pair σ, τ of stopping times taking values in T we have that $E[M_\tau] = E[M_\sigma]$.

For any stopping time τ taking values in $T \cup \{\infty\}$ and any process M the stopped process M^τ

$$M_t^\tau := M_{\tau \wedge t}$$

for $t \in T$ is well defined. If M is a (sub)martingale, M^τ is a (sub)martingale, too, with respect to $(\mathcal{F}_t)_{t \in T}$ and with respect to the stopped filtration $(\mathcal{F}_{t \wedge \tau})_{t \in T}$.

For two stopping times σ, τ taking values in T, and for any martingale M we obtain

$$\mathbb{E}\left[M_\tau \middle| \mathcal{F}_\sigma \right] = M_{\sigma \wedge \tau} \,.$$

For this property we just use that $(M_t - M_{t \wedge \tau})_{t \in T}$ is a martingale with respect to $(\mathcal{F}_t)_{t \in T}$ as difference of two martingales. Stopping this martingale with σ leads to a martingale with respect to $(\mathcal{F}_{t \wedge \sigma})_{t \in T}$, whose evaluation at $\sup T$ and conditional expectation on $\mathcal{F}_\sigma$ leads to the desired property.

Levy-Bernstein inequalities

These basic considerations are already sufficient to derive fundamental inequalities, from which the important set of maximal inequalities follows.

Theorem

Let X be a submartingale on a finite index set T with maximum $supT \in T$, then for any $r \geq 0$ we have

$$rP[\sup_{t\in T} X_t \geq r] \leq E[X_{\sup T} 1_{\{\sup_{t\in T} X_t \geq r\}}] \leq E[X^+_{\sup T}]$$

and

$$rP[\sup_{t\in T} |X_t| \geq r] \leq 3 \sup_{t\in T} E[|X_t|]\,.$$

Proof

Consider the stopping time $\tau = \inf\{s \in T \mid X_s \geq r\}$ (which can take the value ∞) and the predictable strategy

$$V_t := 1_{\{\tau \leq \sup T\}} 1_{\{\tau < t \leq \sup T\}} ,$$

then $E[(V \bullet X)] \geq 0$ by Proposition 2, hence the first assertion follows. For the second one take the sub-martingale $|X|$.

Maximal Inqualities

Theorem

Let M be a martingale on a finite index set T with maximum $\sup T$

$$E[\sup_{t\in T}|M_t|^p] \leq \left(\frac{p}{p-1}\right)^p E[|M_{\sup T}|^p]$$

for all $p > 1$.

Proof

We apply the Levy-Bernstein inequalities from Theorem 3 to the sub-martingale $|M|$. This yields by Fubini's theorem and the Hölder inequality

$$\begin{aligned} E[\sup_{t\in T}|M_t|^p] &= p\int_0^\infty P[\sup_{t\in T}|M_t| > r]r^{p-1}dr \\ &\leq p\int_0^\infty E[|M_{\sup T}|1_{\{\sup_{t\in T}|M_t|\geq r\}}]r^{p-2}dr \\ &= pE\left[|M_{\sup T}|\int_0^{\sup_{t\in T}|M_t|} r^{p-2}dr\right] \\ &= \left(\frac{p}{p-1}\right)E\left[|M_{\sup T}|\sup_{t\in T}|M_t|^{p-1}\right] \\ &\leq \left(\frac{p}{p-1}\right)\|M_{\sup T}\|_p\|\sup_{t\in T}|M_t|^{p-1}\|_q \,, \end{aligned}$$

which yields the result.

Upcrossing inequality

Finally we prove Doob's upcrossing inequality, which is the heart of all further convergence theorems, upwards or downwards. Consider an interval $[a, b]$ and consider a submartingale X, then we denote by $N([a, b], X)$ the number of upcrossings of a trajectory of X from below a to above b. We have the following fundamental lemma:

Lemma

Let X be s submartingale, then

$$E[N([a, b], X)] \leq \frac{E[(X_{\sup T} - a)]^+}{b - a}$$

Proof

Denote $\tau_0 = \min T$, then define recursively the stopping times

$$\sigma_k := \inf\{t \geq \tau_{k-1} \mid X_t \leq a\}$$

and

$$\tau_k := \inf\{t \geq \sigma_k \mid X_t \geq b\}$$

for $k \geq 1$. The process

$$V_t := \sum_{k \geq 1} 1_{\{\sigma_k < t \leq \tau_k\}}$$

is predictable and $V \geq 0$ such as $1 - V$. We conclude by

$$\begin{aligned}(b-a)E[N([a,b],X)] &\leq E[(V \bullet (X-a)^+)_{\sup T}] \\ &\leq E[(1 \bullet (X-a)^+)_{\sup T}] \leq E[(X_{\sup T} - a)^+] \,.\end{aligned}$$

Mind the slight difference of these inequalities to the one in the introduction.

Martingale Convergence

This remarkable lemma allows to prove the following deep convergence results by passing to countable index sets:

Theorem

Let X be an L^1 bounded submartingale on a countable index set T, then there is a set A with probability one such that X_t converges along any increasing or decreasing sequence in T.

Proof

By $\sup_{t \in T} E[|X_t|] < \infty$ we conclude by the Lévy-Bernstein inequality that the measurable random variable $\sup_{t \in T} |X_t|$ is finitely valued, hence along every subsequence there is a finite inferior or superior limit. By monotone convergence we know that for any interval the number of upcrossings is finite almost surely. Consider now A, the intersection of sets with probability one, where the number of upcrossings is finite over intervals with rational endpoints. A has again probability one and on A the process X converges along any increasing or decreasing subsequence, since along monotone sequences a finite number of upcrossings leads to equal inferior and superior limits. Notice that we work here with monotone convergence, since the number of upcrossings for increasing index sets is increasing, however, its expectation is bounded.

Theorem

For any martingale M on any index set we have the following equivalence:

1. *M is uniformly integrable.*
2. *M is closeable at $\sup T$.*
3. *M is L^1 convergent at $\sup T$.*

Proof

If M is closeable on an arbitrary index set T, then by definition there is $\xi \in L^1(\Omega)$ such that $M_t = E[\xi \mid \mathcal{F}_t]$ for $t \in T$, hence

$$E[M_t 1_A] \leq E[E[|\xi| \mid \mathcal{F}_t] 1_A] = E[|\xi| E[1_A \mid \mathcal{F}_t]]$$

for any $A \in \mathcal{F}$, which tends to zero if $P(A) \to 0$, uniformly in t, hence uniform integrability. On the other hand a uniformly integrable martingale is bounded in L^1 and therefore we have one and the same almost sure limit along any subsequence increasing to $\sup T$. If M is uniformly integrable, an almost sure limit is in fact L^1.

Finally assume $M_t \to \xi$ for $t \to \sup T$ in L^1, hence $M_s \to E[\xi \mid \mathcal{F}_s]$, for $s \in T$ and the martingale property, hence $M_s = E[\xi \mid \mathcal{F}_s]$ for any $s \in T$, which concludes the proof.

We obtain the following beautiful corollary:

Corollary

Let M be a martingale on an arbitrary index set and assume $p > 1$, then M_t converges in L^p for $t \to \sup T$ if and only if it is L^p bounded.

Proof

If M is L^p bounded, then it is uniformly integrable (by Doob's maximal inequalities from Theorem 4) and convergence takes place in L^1 by the previous theorem, which in turn by L^p-boundedness is also a convergence in L^p. On the other hand, if M converges in L^p, then it is by the Jensen's inequality also L^p bounded.

Finally we may conclude the following two sided version of closedness:

Theorem

Let T be a countable index set unbounded above and below, then for any $\xi \in L^1$ we have that

$$E[\xi \mid \mathcal{F}_t] \to E[\xi \mid \mathcal{F}_{\pm\infty}]$$

for $t \to \pm\infty$.

Proof

By L^1 boundedness we obtain convergence along any increasing or decreasing subsequence towards limits $M_{\pm\infty}$. The upwards version follows from the previous Theorem 7, the downwards version follows immediately.

Martingale inequalities on uncountable index sets can often be derived from inequalities for the case of countable index sets if certain path properties are guaranteed. From martingale convergence results on countable index sets we can conclude the existence of RCLL versions for processes like martingales, which is the main result of this section.

Submartingale convergence

We need an auxiliary lemma on reverse submartingales first. Of course similar statements hold for supermartingales.

Lemma

Let X be a submartingale on $\mathbb{Z}_{\leq 0}$. Then X is uniformly integrable if and only if $E[X]$ is a bounded (from below) sequence.

Proof

Let $E[X]$ be bounded from below. We can then introduce a Doob-Meyer type decomposition, i.e.

$$A_n := \sum_{k<n} E[X_{k+1} - X_k \mid \mathcal{F}_k] \,,$$

which is well defined since all summands are positive due to submartingality and

$$E[A_0] \leq E[X_0] - \inf_{n \geq 0} E[X_n] < \infty \,.$$

Whence $X = M + A$, where M is a martingale. Since A is uniformly integrable and M is a martingale being closed at 0 by martingale convergence, hence uniformly integrable, also the sum is uniformly integrable. The other direction follows immediately since $E[X_n]$ is decreasing for $n \to \infty$. If it were unbounded from below, it cannot be uniformly integrable.

Regularization

From this statement we can conclude by martingale convergence the following fundamental regularization result:

Theorem

For any submartingale X on $\mathbb{R}_{\geq 0}$ with restriction Y to $\mathbb{Q}_{\geq 0}$ we have:

1. *The process of right hand limits Y^+ exists on $\mathbb{R}_{\geq 0}$ outside some nullset A and $Z := 1_{A^c} Y^+$ is an RCLL submartingale with respect to the augmented filtration $\overline{\mathcal{F}_+}$.*
2. *If the filtration is right continuous, then X has a RCLL version, if and only if $t \mapsto E[X_t]$ is right continuous.*

Proof

The process Y is L^1 bounded on bounded intervals since the positive part is an L^1 bounded submartingale by Jensen's inequality, hence by martingale convergence Theorem 6 we obtain the existence of right and left hand limits and therefore Y^+ is RCLL. Clearly the process Z is adapted to the augmented filtration $\overline{\mathcal{F}_+}$.

Proof

The submartingale property follows readily, too: fix times $s < t$ and choose $s_n \searrow s$ and $t_m \searrow t$, with $s_n < t$ for all $n \geq 1$. Then – by assumption – $E[Y_{t_m} \mid \mathcal{F}_{s_n}] \geq Y_{s_n}$. By martingale convergence to the left we obtain

$$E[Y_{t_m} \mid \mathcal{F}_{s+}] \geq Z_s$$

almost surely. Since the submartingale $(Y_{t_m}))_{m \geq 1}$ has bounded expectations, we conclude L^1-convergence (due to uniform integrability by the previous lemma) and therefore arrive at

$$E[Z_t \mid \overline{\mathcal{F}_{s+}}] \geq Z_s \,.$$

For the second assertion observe that if X is RCLL, then the curve $E[X]$ is right continuous by uniform integrability along decreasing subsequences and the previous lemma. On the other hand if $E[X]$ is right continuous $Z_t = E[Z_t \mid \mathcal{F}_t] \geq X_t$ by limits from the right, but $E[Z_t - X_t] = 0$ by right continuity of $E[X]$, hence Z and X are indistinguishable.

Stochastic Integration

We shall consider now $T = \mathbb{R}_{\geq 0}$ except otherwise mentioned.

Let us introduce some notation: we denote by $\mathbb{S}$ the set of simple predictable processes, i.e. for $\omega \in \Omega$, $s \in T$

$$H_s(\omega) = H_0(\omega)1_{\{0\}}(s) + \sum_{i=1}^{n} H_i(\omega)1_{]T_i(\omega), T_{i+1}(\omega)]}(s)$$

for an increasing, finite sequence of stopping times $0 = T_0 \leq T_1 \leq \ldots T_{n+1} < \infty$ and H_i being $\mathcal{F}_{T_i}$ measurable, by $\mathbb{L}$ the set of adapted, caglad processes and by $\mathbb{D}$ the set of adapted, cadlag processes on $\mathbb{R}_{\geq 0}$.

These vector spaces are endowed with the metric

$$d(X,Y) := \sum_{n \geq 0} \frac{1}{2^n} E\big[|(X-Y)|_n^* \wedge 1\big],$$

which makes $\mathbb{L}$ and $\mathbb{D}$ complete topological vector spaces. We call this topology the ucp-topology ("uniform convergence on compacts in probability"). Notice that predictable strategies as well as integrators are considered $\mathbb{R}$ valued here, which, however, *contains* the $\mathbb{R}^n$ case.

Notice that we are dealing here with topological vector spaces, which are not even locally convex. This leads also to the phenomenon that the metric does not detect boundedness of sets, which is defined in the following way: A subset B of a topological vector space $\mathbb{D}$ is called bounded, if it can be absorbed by any open neighborhood U of zero, i.e. there is $R > 0$ such that $B \subset R\, U$. For instance for the space of random variables $L^0(\Omega)$ this translates to the following equivalent statement: a set B of random variables is bounded in probability if for every $\epsilon > 0$ there is $c > 0$ such that

$$P[|Y| \geq c] < \epsilon$$

for $Y \in B$, which is of course not detectable by the metric.

Good integrators

Definition

An adapted, cadlag process X is called good integrator if the map

$$J_X : \mathbb{S} \to \mathbb{D}$$

with

$$(H \bullet X)_t := J_X(H)_t := H_0 X_0 + \sum_{i=1}^{n} H_i (X_{T_{i+1} \wedge t} - X_{T_i \wedge t}),$$

for $H \in \mathbb{S}$, is continuous with respect to the ucp-topologies.

Remark

It would already be sufficient to require a good integrator to satisfy the following property: for every stopped process X^t, $t \geq 0$, the map $I_{X^t} : \mathbb{S}_u \to L^0(\Omega)$, where $I_{X^t}(H) := J_X(H)_\infty$, from uniformly bounded, simple predictable processes *with the uniform topology* $\mathbb{S}_u$ to random variables with convergence in probability, is continuous, i.e.

$$I_{X^t}(H^k) := H_0^k X_0 + \sum_{i=1}^{n} H_i^k (X_{T_{i+1} \wedge t} - X_{T_i \wedge t}) \to 0 \qquad (2)$$

if $H^k \to 0$ uniformly on $\Omega \times \mathbb{R}_{\geq 0}$.

Remark

From continuity with respect to the uniform topology continuity with respect to the ucp topologies, as claimed in the definition of a good integrator, immediately follows.

Indeed assume that I_X is continuous with respect to the uniform topology, fix $n \geq 0$ and a sequence $H^k \to 0$, which tends to 0 uniformly, then choose $c \geq 0$ and define a sequence of stopping times

$$\tau^k := \inf\{t \mid |(H^k \bullet X)_t| \geq c\}$$

for $k \geq 0$, then

$$P[|(H^k \bullet X)|_n^* \geq c] = P[|(H^k 1_{[0,\tau^k]} \bullet X)_n| \geq c] =$$
$$= P[|(H^k 1_{[0,\tau^k]} \bullet X^n)_\infty| \geq c] \to 0$$

as $k \to \infty$ by assumption, hence $J_X : \mathbb{S}_u \to \mathbb{D}$ is continuous, i.e. we can map to processes without loosing continuity.

Remark

Take now a sequence $H^k \to 0$ in ucp, and choose $c \geq 0$, $\epsilon > 0$ and $n \geq 0$. Then there is some $\eta > 0$ such that

$$P[|(H \bullet X)|_n^* \geq c] \leq \epsilon$$

for $\|H\|_\infty \leq \eta$ by continuity of $J_X : \mathbb{S}_u \to \mathbb{D}$. Define furthermore stopping times

$$\rho^k := \inf\{s \mid |H_s^k| > \eta\}$$

then we obtain

$$P[|(H^k \bullet X)|_n^* \geq c] \leq P[|(H^k 1_{[0,\rho^k]} \bullet X)_n| \geq c] + P[\rho^k < n] < 2\epsilon$$

if k is large enough since $P[\rho^k < n] \to 0$ as $k \to \infty$.

Localization

Clearly Property (2) holds if it only holds locally: indeed let τ^n be a localizing sequence, i.e. $\tau^n \nearrow \infty$ with X^{τ^n} being a good integrator. Fix $t \geq 0$ and a sequence $H^k \to 0$, which tends to 0 uniformly, then

$$P[|I_{X^t}(H^k)| \geq c] \leq P[|I_{X^{\tau^n \wedge t}}(H^k)| \geq c] + P[\tau^n \leq t]$$

for every $n \geq 0$. Hence we can choose n large enough, such that the second term is small by the localizing property, and obtain for k large enough that the first term is small by X^{τ^n} being a good integrator.

Density of simple processes

The set $\mathbb{S}$ is dense in the ucp-topology in $\mathbb{L}$, even the bounded simple predictable processes are dense.

Finite variation processes are good integrators

Any process of finite variation A is a good integrator, since for every simple (not even predictable) process it holds that

$$|J_A(H)_t| = |\int_0^t H_s dA_s| \leq \|H\|_\infty \Big(\int_0^t d|A|_s + |A_0| \Big)$$

almost surely, for $t \geq 0$.

Ito's insight

By Ito's fundamental insight square integrable martingales M are good integrators, since

$$E\left[(H \bullet M)_t{}^2\right] \leq \|H\|_\infty^2 E[|M_t|^2]$$

holds true for simple, bounded and predictable processes $H \in \mathbb{S}_u$.

Burkholder inequality

By the following elementary inequality due to Burkholder we can conclude that martingales are good integrators: for every martingale M and every simple, bounded and predictable process $H \in \mathbb{S}_u$ it holds that

$$cP(|(H \bullet M)|_1^* \geq c) \leq 18 \|H\|_\infty \|M_1\|_1$$

for all $c \geq 0$.

Since the inequality is crucial for our treatment, we shall prove it here, too. Notice that we are just dealing with integrals with respect to simple integrands, hence we can prove it for discrete martingales on a finite set of time points.

Proof

Let M be a non-negative martingale first and H bounded predictable with $\|H\|_\infty \leq 1$, then $Z := M \wedge c$ is a supermartingale and we have

$$cP(|(H \bullet M)|_1^* \geq c) \leq cP(|M|_1^* \geq c) + cP(|(H \bullet Z)|_1^* \geq c)\,.$$

Since Z is a super-martingale we obtain by the Doob-Meyer decomposition for discrete super-martingales $Z = \tilde{M} - A$ that

$$|(H \bullet Z)| \leq |(H \bullet \tilde{M})| + A\,,$$

i.e. we have an upper bound being a sub-martingale.

Proof

With $|(H \bullet \tilde{M})| + A$ also its square is a sub-martingale. Hence we can conclude by Lemma 3 that

$$cP(|(H \bullet M)|_1^* \geq c) \leq E[M_1] + 2\frac{1}{c}E[(H \bullet \tilde{M})_1^2 + A_1^2]\,,$$

since

$$\begin{aligned} cP(|(H \bullet Z)|_1^* \geq c) \leq cP\Big(\big||(H \bullet \tilde{M})| + A\big|_1^* \geq c) \leq \\ \leq \frac{1}{c}E\left[\left(|(H \bullet \tilde{M})|_1 + A_1\right)^2\right] \leq 2\frac{1}{c}E[(H \bullet \tilde{M})_1^2 + A_1^2]\,. \end{aligned}$$

Proof

Ito's insight allows to estimate the variance of the stochastic integral at time 1 by $E[\tilde{M}_1^2]$. Both quantities $\tilde{M}$ and A of the Doob-Meyer decomposition may, however, be estimated through $E[A_1^2] \leq E[\tilde{M}_1^2] \leq 2cE[Z_0] \leq 2cE[M_0]$, see (1), since Z is non-negative (so $A \leq \tilde{M}$ holds true) and $Z \leq c$. This leads to an upper bound

$$cP(|(H \bullet M)|_1^* \geq c) \leq 9E[M_0] \,.$$

Writing a martingale as difference of two non-negative martingales leads to the desired result.

Martingales are good integrators

Apparently the result translates directly to the fact that M is a good integrator. We actually immediately obtain that $J_X : \mathbb{S}_u \to \mathbb{D}$ is continuous, wherefrom – as we have seen before – the continuity even with respect to the ucp topology on $\mathbb{S}$ follows.

Associativity et al.

By density and continuity we can extend the map J_X to all caglad processes $Y \in \mathbb{L}$, which defines the stochastic integral $(Y \bullet X)$. As a simple corollary we can prove the following proposition:

Proposition

Let $H, G \in \mathbb{L}$ be given and let X be a good integrator, then $(G \bullet X)$ is a good integrator and $(H \bullet (G \bullet X)) = (HG \bullet X)$.

Proof

Let X be a good integrator, then $J_X : \mathbb{L} \to \mathbb{D}$ is continuous with respect to the ucp topologies. Let (H_k) be a sequence in $\mathbb{S}_u$ converging uniformly to 0, then $H_k G$ converges ucp to 0 for every $G \in \mathbb{L}$, whence $J_{(G \bullet X)}$ which satisfies

$$J_{(G \bullet X)}(H) = (HG \bullet X)$$

is obviously continuous, and the desired formula holds by continuous extension.

Approximation results

We know that $H \mapsto (H \bullet X)$ is continuous with respect to the ucp topologies on the left and right hand side. However, actually a bit more is true.

Most important for the calculation and understanding of stochastic integrals is the following approximation result: a *sequence of partition tending to identity* Π^k consists of stopping times $0 = T_0^k \leq \ldots \leq T_{i_k}^k < \infty$ with mesh $\sup_i(T_{i+1}^k - T_i^k)$ tending to 0 and $\sup_i T_i^k \to \infty$. We call the sequence of cadlag processes

$$Y^{\Pi^k} := \sum_i Y_{T_i^k} 1_{[T_i^k, T_{i+1}^k[}$$

a sampling sequence for a cadlag process Y along Π^k, for $k \geq 0$. Notice that we do not necessarily have that $Y^{\Pi^k} \to Y$ in ucp, nor $Y_-^{\Pi^k} \to Y_-$ due to the presence of large jumps.

Example

One important sequence of partition is constructed by a truncation of the following one: let $Y \in \mathbb{D}$ be a cadlag process. For $n \geq 0$ we can define a double sequence of stopping times τ_i^n

$$\tau_0^n := 0 \text{ and } \tau_{i+1}^n := \inf\{s \geq \tau_i^n \mid |Y_s - Y_{\tau_i^n}| \geq \frac{1}{2^n}\}$$

for $i \geq 0$. This defines a sequence of partitions

$$\Pi^n = \{\tau_0^n \leq \ldots \tau_{n2^n}^n\}$$

tending to identity. We have that

$$|Y_- - Y_-^{\Pi^n}| \leq \frac{1}{2^n},$$

hence $Y_-^{\Pi^n} \to Y_-$ in the ucp topology, and also $Y^{\Pi^n} \to Y$ in ucp.

Theorem

For any good integrator X we obtain that

$$(Y_-^{\Pi^k} \bullet X) \to (Y_- \bullet X)$$

in the ucp topology in general (even though Y^{Π^k} does not necessarily converge to Y in ucp), as well as the less usually stated but equally true ucp convergence result

$$(Y_-^{\Pi^k} \bullet X^{\Pi^k}) \to (Y_- \bullet X)_- \,.$$

Remark

Notice that so far we have no understanding on the continuity of $X \mapsto (H \bullet X)$ on the set of good integrators. Of course it is not ucp continuous. The second assertion gives an answer in a very specific case.

Proof

We know by previous remarks that there are sequences Y^l of simple cadlag processes converging ucp to Y, where $Y^l_- \to Y_-$ holds true for the associated left continuous processes. Hence we can write

$$((Y_- - Y_-^{\Pi^k})\bullet X) = ((Y_- - Y_-^l)\bullet X) + ((Y_-^l - (Y^l)_-^{\Pi^k})\bullet X) + (((Y^l)_-^{\Pi^k} - Y_-^{\Pi^k})\bullet X)$$

where the first and third term converge of course in ucp as $l \to \infty$, the third even uniformly in k. The middle term is seen to converge by direct inspection.

Proof

The proof of the second assertion follows from the fact that

$$(Y_-^{\Pi^k} \bullet X^{\Pi^k}) = (Y_- \bullet X^{\Pi^k}) \to (Y_- \bullet X)_- \,,$$

where the limit assertion follows from the fact that $(Y_- \bullet X^{\Pi^k})$ only differs from $(Y_-^{\Pi^k} \bullet X)$ on the 'last' interval before t is reached by the stopping times in the partition Π^k, which is a quantity converging to 0 plus the last jump, i.e.

$$Y_{T_i}(X_{T^k_{i+1} \wedge t} - X_{T^k_i \wedge t}) 1_{[T_{i_t}, T_{i_t+1}[}(t) - Y_t \Delta X_t \to 0$$

in ucp. i_t is chosen such that $t \in [T_{i_t}, T_{i_t+1}[$.

Definition

Let X, Y be good integrators, then we define the quadratic (co-) variation process by

$$[X, Y] := XY - (X_- \bullet Y) - (Y_- \bullet X)\,.$$

Remark

Quadratic variation $[X, X]$ is an non-decreasing (hence finite variation) process for any good integrator. As a consequence of the previous approximation theorem we obtain of course for two good integrators X, Y that

$$[X^{\Pi^k}, Y^{\Pi^k}]_t = \sum_{T^k_{i+1} < t} (X_{T^k_{i+1}} - X_{T^k_i})(Y_{T^k_{i+1}} - Y_{T^k_i}) \to [X, Y]_{t-} \,,$$

in ucp, since quadratic co-variation can be expressed by stochastic integrals as given in the definition. Whence by adding a last jump

$$\sum_i (X_{T^k_{i+1} \wedge t} - X_{T^k_i \wedge t})(Y_{T^k_{i+1} \wedge t} - Y_{T^k_i \wedge t}) \to [X, Y]_t$$

again in ucp.

Convention

Let us fix an important notations here: we shall always assume that $X_{0-} = 0$ (a left limit coming from negative times), whereas X_0 can be different from zero, whence ΔX_0 is not necessarily vanishing.

Proposition

Let $H \in \mathbb{L}$ be fixed, as well as two good integrators X, Y. Then

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

Proof.

Let $H \in \mathbb{S}$ be fixed. Then apparently

$$[(H \bullet X), Y] = \left(H \bullet [X, Y]\right),$$

whence by continuity with respect to ucp topologies also the left hand side is continuous with respect to ucp topologies, which proves the result by continuity of $H \mapsto (H \bullet X)$ with respect to ucp. □

Ito's theorem

The set of good integrators is a vector space, in fact even an algebra. More precisely: given finitely many good integrators $X^1, \ldots, X^n$ then $f(X^1, \ldots, X^n)$ is also a good integrator for any C^2 function f.

It is remarkable that Ito's theorem can be concluded from its version for piece-wise constant processes due to the following continuity lemma, which complements results which have already been established for the approximation of stochastic integrals.

We state an additional continuity lemma:

Lemma

Let $X^1, \ldots, X^n$ be good integrators, Π^k a sequence of partitions tending to the identity and $f : \mathbb{R}^n \to \mathbb{R}$ a C^2 function, then for $t \geq 0$

$$
\begin{aligned}
&\sum_{s \leq t} \Big\{ f(X_s^{\Pi^k}) - f(X_{s-}^{\Pi^k}) - \sum_{i=1}^n \partial_i f(X_{s-}^{\Pi^k}) \Delta X_s^{i,\Pi^k} - \\
&- \frac{1}{2} \sum_{i,j=1}^n \partial_{ij}^2 f(X_{s-}^{\Pi^k}) \Delta X_s^{i,\Pi^k} \Delta X_s^{j,\Pi^k} \Big\} \\
&\to_{k \to \infty} \sum_{s \leq t} \Big\{ f(X_s) - f(X_{s-}) - \sum_{i=1}^n \partial_i f(X_{s-}) \Delta X_s^i - \\
&- \frac{1}{2} \sum_{i,j=1}^n \partial_{ij}^2 f(X_{s-}) \Delta X_s^i \Delta X_s^j \Big\} \,,
\end{aligned}
$$

where the limit can be even understood in ucp topology.

Remark

Here we mean with X^{Π^k} the cadlag version of the sampled process as introduced before.

Proof

The proof of this elementary lemma relies on Taylor expansion of f: apparently the finitely many summands of the approximating series are small at s if $(\Delta X_s^{\Pi^k})^{(2+)}$ is small, hence only those jumps remain after the limit, which are at time points where X actually jumps. Let us make this precise: first we know – by the very existence of quadratic variation – that

$$\sum_{s \leq t} (\Delta X^i)^2 \leq [X^i, X^i]_t < \infty$$

almost surely. Fix $t \geq 0$ and $\epsilon > 0$, then we find for every $\omega \in \Omega$ a finite set A_ω of times up to t, where X jumps in a large way (defined by the condition on B), and a possibly countable set of times B_ω up to t, where X jumps *and* $\sum_{s \in B} \|\Delta X\|^2 \leq \epsilon^2$, since every cadlag path has at most countably many jumps up to time t.

Proof

Furthermore we know that

$$f(y) - f(x) - \sum_{i=1}^{n} \partial_i f(x)(y-x)^i - \frac{1}{2}\sum_{i,j=1}^{n} \partial_{ij}^2 f(x)(y-x)^i (y-x)^j =$$
$$= o\big(\|y-x\|\big)\|y-x\|^2$$

as $y \to x$. This means that for $\omega \in \Omega$ we can split the approximating sum into two sums denoted by $\sum_A$, where $]T_i^k(\omega), T_{i+1}^k(\omega)] \cap A_\omega \neq \emptyset$, and $\sum_B$ corresponding to jumps which appear at B and $\sum_C$ over intervals, where no jumps appear in the limit. We then obtain an estimate for the limiting sum $\sum_B$ of the type

$$\sum_B \leq 2\epsilon^2 o\big(\epsilon\big)$$

for k large enough by uniform continuity of continuous functions on compact intervals and jump size at most ϵ. Furthermore we obtain

$$\sum_C \leq \|[X^{T^k}, X^{T^k}]\| o\big(\max_i 1_{]T_i^k, T_{i+1}^k] \cap (A \cup B) = \emptyset} \|X_{T_{i+1}^k} - X_{T_i^k}\|\big)\,.$$

Ito's formula

We are now able to prove Ito's formula in all generality:

Theorem

Let $X^1, \dots, X^n$ be good integrators and $f : \mathbb{R}^n \to \mathbb{R}$ a C^2 function, then for $t \geq 0$

$$f(X_t) = \sum_{i=1}^{n} (\partial_i f(X_-) \bullet X^i)_t + \frac{1}{2} \sum_{i,j=1}^{n} (\partial^2_{ij} f(X_-) \bullet [X^i, X^j])_t +$$

$$+ \sum_{0 \leq s \leq t} \left\{ f(X_s) - f(X_{s-}) - \sum_{i=1}^{n} \partial_i f(X_{s-}) \Delta X^i_s - \frac{1}{2} \sum_{i,j=1}^{n} \partial^2_{ij} f(X_{s-}) \Delta X^i_s \Delta X^j_s \right\}.$$

Notice that $f(X_0)$ is in the second sum since we agreed that $f(X_{0-}) = 0$.

Proof

Let Π^k be a sequence of partitions tending to the identity, then Ito's formula reads by careful inspection

$$f(X_t^{\Pi^k}) = \sum_{i=1}^{n} (\partial_i f(X_-^{\Pi^k}) \bullet X^{i,\Pi^k})_t + \frac{1}{2} \sum_{i,j=1}^{n} (\partial_{ij}^2 f(X_-^{\Pi^k}) \bullet [X^{i,\Pi^k}, X^{j,\Pi^k}])_t +$$

$$+ \sum_{0<s\leq t} \Big\{ f(X_s^{\Pi^k}) - f(X_{s-}^{\Pi^k}) - \sum_{i=1}^{n} \partial_i f(X_{s-}^{\Pi^k}) \Delta X_s^{i,\Pi^k} -$$

$$- \frac{1}{2} \sum_{i,j=1}^{n} \partial_{ij}^2 f(X_{s-}^{\Pi^k}) \Delta X_s^{i,\Pi^k} \Delta X_s^{j,\Pi^k} \Big\},$$

since the process is piece-wise constant and the sum is just telescoping. By the previously stated convergence result, however, this translates directly – even in ucp convergence – to the limit for $k \to \infty$, which is Ito's formula.

Quadratic Pure jump good integrators

We call a good integrator *quadratic pure jump* if $[X,X]_t = \sum_{s\leq t} (\Delta X_s)^2$ for $t \geq 0$. It follows from Ito's formula that every cadlag, adapted and finite variation process X is quadratic pure jump.

Indeed the finite variation property yields a well-known Ito formula for $f(x) = x^2$ (notice that the second order terms is missing in the sum) of the type

$$X_t^2 = 2(X_- \bullet X) + \sum_{s\leq t} \{X_s^2 - X_{s-}^2 - 2X_{s-}\Delta X_s\} = 2(X_- \bullet X) + \sum_{s\leq t} (\Delta X_s)^2 \,,$$

which yields the result on the quadratic variation. Hence for every good integrator M we obtain

$$[X,M]_t = \sum_{s\leq t} \Delta X \Delta M$$

for finite variation processes X with complete analogous arguments.

Stochastic exponentials

An instructive example how to calculate with jump processes is given by the following process: let X be a good integrator with $X_0 = 0$, then the process

$$Z_t = \exp\big(X_t - \frac{1}{2}[X,X]_t\big) \prod_{0 \le s \le t} (1 + \Delta X_s) \exp\big(- \Delta X_s + \frac{1}{2}(\Delta X_s)^2\big)$$

satisfies $Z_t = 1 + (Z_- \bullet X)_t$ and is called stochastic exponential.

Proof

For the proof we have to check that the infinite product is actually converging and defining a good integrator. We show this by proving that it defines an adapted, cadlag process of finite variation. We only have to check this for jumps smaller than $\frac{1}{2}$, i.e. we have to check whether

$$\sum_{s\leq t}\{\log(1+U_s)-U_s+\frac{1}{2}U_s^2\}$$

converges absolutely, where $U_s := \Delta X_s 1_{\{|\Delta X_s|\leq\frac{1}{2}\}}$, for $s\geq 0$. This, however, is true since $|\log(1+x)-x+\frac{1}{2}x^2|\leq Cx^3$ for $|x|\leq\frac{1}{2}$ and $\sum_{s\leq t}{\Delta X_s}^2\leq[X,X]<\infty$ almost surely.

Proof

Hence we can apply Ito's formula for the function $\exp(x_1)x_2$ with good integrators

$$X_t^1 = X_t - \frac{1}{2}[X, X]_t$$

and

$$X_t^2 = \prod_{0 \le s \le t} (1 + \Delta X_s) \exp\big(- \Delta X_s + \frac{1}{2}(\Delta X_s)^2 \big) .$$

Proof

This leads to a decomposition for Z_t like

$$1 + (Z_- \bullet X)_t - \frac{1}{2}(Z_- \bullet [X, X])_t + (\exp(X^1_-) \bullet X^2)_t + \frac{1}{2}(Z_- \bullet [X^1, X^1])_t +$$

$$+ \sum_{s \le t} \{ Z_s - Z_{s-} - Z_{s-}\Delta X^1_s - \exp(X^1_{s-})\Delta X^2_s - \frac{1}{2} Z_{s-}(\Delta X^1_s)^2 \} =$$

$$= 1 + (Z_- \bullet X)_t - \frac{1}{2}(Z_- \bullet [X, X])_t + \frac{1}{2}(Z_- \bullet [X^1, X^1])_t +$$

$$+ \sum_{s \le t} \{ Z_{s-}\Delta X_s - Z_{s-}\Delta X^1_s - \frac{1}{2} Z_{s-}(\Delta X^1_s)^2 \} =$$

$$= 1 + (Z_- \bullet X)_t - \frac{1}{2}(Z_- \bullet [X, X])_t + \frac{1}{2}(Z_- \bullet [X^1, X^1])_t +$$

$$+ \sum_{s \le t} \{ \frac{1}{2} Z_{s-}(\Delta X_s)^2 - \frac{1}{2} Z_{s-}(\Delta X^1_s)^2 \} =$$

$$= 1 + (Z_- \bullet X)_t \,,$$

Proof

since

$$Z_s = Z_{s-} \exp\big(\Delta X_s - \frac{1}{2}(\Delta X_s)^2\big)(1 + \Delta X_s)\exp\big(-\Delta X_s + \frac{1}{2}(\Delta X_s)^2\big)$$

holds true, for $s \geq 0$.

Lévy's Theorem

Another remarkable application is Lévy's theorem: consider local martingales $B^1, \ldots, B^n$ starting at 0 with continuous trajectories such that $[B^i, B^j]_t = \delta^{ij} t$ for $t \geq 0$. Then $B^1, \ldots, B^n$ are standard Brownian motions.

Martingality of some stochastic integrals

For this theorem we have to check that stochastic integrals along locally square integrable martingales are locally square integrable martingales: indeed let X be a locally square-integrable martingale and $H \in \mathbb{L}$, then by localizing and the formula

$$(H \bullet X)^{\tau} = (H \bullet X^{\tau}) = (H 1_{]0,\tau]} \bullet X^{\tau}) = (H 1_{]0,\tau]} \bullet X)$$

we can assume that H is bounded and X is in fact a square integrable martingale with $E[X_\infty^2] < \infty$. Then, however, we have by Ito's insight that for sequence of partitions tending to identity the process

$$(H^{\Pi^k} \bullet X)$$

is a square integrable martingale, which satisfies additionally

$$E[(H^{\Pi^k} \bullet X)_\infty^2] \leq \|H\|_\infty^2 E[X_\infty^2] \, .$$

By martingale convergence this means that the limit in probability of the stochastic integrals is also a square integrable martingale, whence the result.

Proof of Lévy's Theorem

This can be readily applied to the stochastic process

$$M_t := \exp\left(\mathrm{i}\langle\lambda, B_t\rangle + \frac{t}{2}\|\lambda\|^2\right)$$

for $t \geq 0$ and $\lambda \in \mathbb{R}^n$, which is a bounded local martingale by the previous consideration and Ito's formula. A bounded local martingale is a martingale, whence

$$E[\exp(\mathrm{i}\langle\lambda, B_t - B_s\rangle) \,|\, \mathcal{F}_s] = \exp(-\frac{t-s}{2}\|\lambda\|^2)\,.$$

Stochastic Integration for predictable integrands

For many purposes (i.e. martingale representation) it is not enough to consider only caglad integrands, but predictable integrands are needed. This cannot be achieved universally for all good integrators, but has to be done case by case. The main tool for this purpose are $\mathcal{H}^p$ spaces, for $1 \leq p < \infty$, which are spaces of martingales with certain integrability properties, the most important being $\mathcal{H}^1$. We present first the $\mathcal{H}^p$ and specialize then to $p = 1$ and $p = 2$.

Main tool for the analysis are the Burkholder-Davis-Gundy inequalities:

Theorem

For every $p \geq 0$ there are constants $0 < c_p < C_p$ such that for every martingale

$$c_p E\left[[M,M]_\infty^{\frac{p}{2}}\right] \leq E\left[(|M|_\infty^*)^p\right] \leq C_p E\left[[M,M]_\infty^{\frac{p}{2}}\right]$$

holds true.

The inequalities follow from the same inequalities for discrete martingales, which can be proved by deterministic methods. In fact equations of the type

$$(h \bullet M)_T + [M, M]_T^{\frac{p}{2}} \leq (|M|_T^*)^p \leq C_p [M, M]_T^{\frac{p}{2}} + (g \bullet M)_T$$

hold true, with predictable integrands h, g and martingales M on a finite index set with upper bound T hold, see works on robust finance.

The case $p = 1$

For discrete martingales we can consider deterministic inequalities of the type

$$\sqrt{[x,x]_N} \leq 3|x|_N^* - (h \bullet x)_N$$

and

$$|x|_N^* \leq 6\sqrt{[x,x]_N} + 2(h \bullet x)_N$$

for a "predictable" strategy

$$h_i := \frac{x_i}{\sqrt{[x,x]_i + (|x|_i^*)^2}}$$

for $i = 1, \ldots, n$. Here we consider sequences $0 = x_0, x_1, \ldots, x_N$.

Proof

Let M be a martingale and let us take a sequence of refining partitions Π^n tending to identity, for which $M^{\Pi^n} \to M$ in ucp and $[M^{\Pi^n}, M^{\Pi^n}] \to [M, M]$ in ucp. Fix some time horizon $T > 0$, then by monotone convergence

$$E\left[(|M^{\Pi^n}|_T^*)^p\right] \to E\left[(|M|_T^*)^p\right]$$

as $n \to \infty$, since the sequence of partitions is refining. If $E\left[(|M|_T^*)^p\right] = \infty$, we obtain that all three quantities are infinity. If $E\left[(|M|_T^*)^p\right] < \infty$ we obtain by dominated convergence that

$$E\left[[M^{\Pi^m} - M^{\Pi^n}, M^{\Pi^m} - M^{\Pi^n}]_T^{\frac{p}{2}}\right] \leq E\left[(|M^{\Pi^m} - M^{\Pi^n}|_T^*)^p\right] \to 0\,.$$

Proof

This means by

$$E\left[\left|[M^{\Pi^m}, M^{\Pi^m}]^{\frac{1}{2}} - [M^{\Pi^n}, M^{\Pi^n}]^{\frac{1}{2}}_T\right|^p\right] \leq E\left[[M^{\Pi^m} - M^{\Pi^n}, M^{\Pi^m} - M^{\Pi^n}]^{\frac{p}{2}}_T\right]$$

the L^p convergence of the quadratic variations to $[M, M]$. This yields the result for any $T > 0$ and hence for $T \to \infty$.

Remark

The case $p = 2$ can be readily derived from Doob's maximal inequality, see Theorem 4, and we obtain

$$E\big[[M,M]_\infty\big] = E\big[M_\infty^2\big] \leq E\big[(|M|_\infty^*)^2\big] \leq 4E[M_\infty^2\big] = E\big[[M,M]_\infty\big]\,.$$

Definition

Let $p \geq 1$ be given. Define the vector space $\mathcal{H}^p$ as set of martingales M where

$$\|M\|_{\mathcal{H}^p}^p := E\left[(|M|_\infty^*)^p\right] < \infty$$

holds true.

BDG inequalities

By the Burkholder-Davis-Gundy (BDG) inequalities the following theorem easily follows immediately:

Theorem

For $p \geq 1$ the space $\mathcal{H}^p$ is a Banach space with equivalent norm

$$M \mapsto E\left[[M,M]_\infty^{\frac{p}{2}}\right]^{\frac{1}{p}}.$$

For $p = 2$ the equivalent norm is in fact coming from a scalar product

$$(M,N) \mapsto E\left[[M,N]_T\right].$$

Additionally we have the following continuity result: $M^n \to M$ in $\mathcal{H}^p$, then $(Y \bullet M_n) \to (Y \bullet M)$ in ucp for any left-continuous process $Y \in \mathbb{L}$. In particular $[M^n, N] \to [M, N]$ in ucp.

In the next step we consider a weaker topology of L^p type on the set of simple predictable integrands. The following lemma tells about the closure with respect to this topology.

Lemma

Let A be an increasing finite variation process and V a predictable process with

$$(|V|^p \bullet A)_t < \infty \,,$$

then there exists a sequence of bounded, simply predictable processes V^n in $b\mathbb{S}$ such that

$$(|V - V^n|^p \bullet A)_t \to 0 \,,$$

as $n \to \infty$.

Proof.

By monotone class arguments it is sufficient to prove the lemma for caglad processes, for which it is, however, clear, since they can be approximated in the ucp topology by simple predictable processes. □

The main line of argument is to construct for predictable processes V satisfying certain integrability conditions with respect to $[M, M]$ a stochastic integral $(V \bullet M)$ for $M \in \mathcal{H}^p$. We take the largest space $\mathcal{H}^1$ in order to stay as general as possible.

Proposition

Let $M \in \mathcal{H}^1$ be fixed and let V^n be a sequence of bounded, simple predictable processes such that

$$E\left[(|V - V^n|^2 \bullet [M, M])_\infty^{\frac{1}{2}}\right] \to 0$$

then the sequence $(V^n \bullet M)$ is a Cauchy sequence in $\mathcal{H}^1$ defining an element $(V \bullet M)$, which does only depend on V and not on the approximating sequence V^n and which is uniquely determined by

$$[(V \bullet M), N] = (V \bullet [M, N])$$

for martingales N.

Proof

This is a direct consequence of the Burkholder-Davis-Gundy inequalities, since

$$E\left[|(V^n \bullet M) - (V^m \bullet M)|^*_\infty)\right] \leq C_1 E\left[((V^n - V^m)^2 \bullet [M,M])^{\frac{1}{2}}_\infty\right] \to 0$$

as $n, m \to \infty$. Whence $(V \bullet M)$ is a well-defined element of $\mathcal{H}^1$, which only depends on V and not on the approximating sequence. For all martingales N and all simple predictable strategies the formula

$$[(V^n \bullet M), N] = (V^n \bullet [M,N])$$

holds true by basic rules for caglad integrands. By passing to the limit we obtain the general result. Uniqueness is clear since $[M,M] = 0$ means $M =$ by Burkholder-Davis-Gundy inequalities.

Definition

Let $M \in \mathcal{H}^1$, then we denote by $L^1(M)$ the set of predictable processes V such that

$$E\left[(|V|^2 \bullet [M, M])_\infty^{\frac{1}{2}}\right] < \infty\,.$$

Apparently we have constructed a bounded linear map $L^1(M) \to \mathcal{H}^1$, $V \mapsto (V \bullet H)$. The set of integrands $L^1(M)$ is not the largest one, we can still generalize it by localization, which defines the set $L(M)$: a predictable process V is said to belong to $L(M)$ if $(|V|^2 \bullet [M,M])_\infty^{\frac{1}{2}}$ is locally integrable, which means for bounded variation processes nothing else than just being finite.

Notice that this is the largest set of integrands given that we require that the integral is a semi-martingal having a quadratic variation, which coincides with $(V^2 \bullet [M,M])$. Notice also that by the same argument every local martingale is in fact locally $\mathcal{H}^1$, which in turn means that we can define for any semi-martingale a largest set of integrands.

The Emery topology

The Emery topology on the set of semimartingales $\mathbb{SEM}$ is defined by the metric

$$d_E(S_1, S_2) := \sup_{K \in \mathbb{S},\, \|K\|_\infty \leq 1} E\left[|(K \bullet (S_1 - S_2))|_1^* \wedge 1\right].$$

We can by means of the Bichteler-Dellacherie theorem easily prove the following important theorem.

Theorem

The set of semi-martingales $\mathbb{SEM}$ is a topological vector space and complete with respect to the Emery topology.

Proof

Obviously d_E defines a metric and a Cauchy sequence $(S_n)_{n\geq 1}$ in d_E is a Cauchy sequence in d, so there is a càdlàg process S which is the pathwise uniform limit of the semi-martingales S_n. We have to show that S is a semi-martingale. We show that by proving that I_S is continuous on $\mathbb{S}_u$ with respect to the uniform topology, which is equivalent to the fact that the set $\{(K \bullet S)_1 \mid K \in \mathbb{S}, \|K\|_\infty \leq 1\}$ is bounded in probability. Fix $1 > \epsilon > 0$, then for $c > 0$

$$\mathbb{P}\big((K \bullet S)_1 \geq c\big) \leq \mathbb{P}\big((K \bullet S - S_n)_1 \geq c\big) + \mathbb{P}\big((K \bullet S_n)_1 \geq c\big) \tag{3}$$

$$\leq d_E(S, S_n) + \mathbb{P}\big((K \bullet S_n)_1 \geq c\big)\,. \tag{4}$$

Now we choose n large enough to make the first term smaller than ϵ. Since S_n is a semi-martingale I_{S_n} is continuous, hence the set $\{(K \bullet S_n)_1 \mid K \in \mathbb{S}, \|K\|_\infty \leq 1\}$ is bounded in probability, which in turn means that we can choose c large enough such that the second term is smaller than ϵ. Hence both terms are small and therefore S is a good integrator.

Theorem

For every semi-martingale S the map I_S defined on $\mathbb{S}$ extends to a continuous map J_M from the space $\mathbb{L}$ of càglàd processes to $\mathbb{SEM}$ of semi-martingales.

Proof

It is sufficient to show the result for martingales S, since the rest follows by localization and the respective theorem for finite variation processes. Let S be a martingale. Take a sequence of simple predictable processes H_n which converges ucp to 0, i.e. $\mathbb{P}(|H_n|_1^* \geq b) \to 0$ as $n \to \infty$. Fix furthermore $K \in \mathbb{S}$ with $\|K\|_\infty \leq 1$. We can decompose $H_n = H_n' + H_n''$ where $H_n' := H_n 1_{\{|H_n|^* \geq b\}}$ for some $b \geq 0$. This decomposition is of course done in $\mathbb{S}$. Observe that $H_n' H_n'' = 0$ for all $n \geq 1$. Now we can estimate through

$$\{|(KH_n \bullet S)|_1^* \geq c\} \subset \{|H_n|_1^* \geq b\} \cup \{|(KH_n'' \bullet S)|_1^* \geq c\}$$

the probabilities directly

$$c\mathbb{P}(|(KH_n \bullet S)|_1^* \geq c) \leq c\mathbb{P}(|H_n|_1^* \geq b) + 18\|KH_n''\|_\infty \|S_1\|_1 \,,$$

where we notice that $\|H_n''\|_\infty \leq b$ and $\mathbb{P}(|H_n|_1{}^* \geq b) \to 0$ as $n \to \infty$.

Proof

This, however, yields that

$$\sup_{K \in \mathbb{S},\, \|K\|_\infty \le 1} E\big[|(KH_n \bullet S)|_1^* \wedge 1\big] \le \mathbb{P}(|H_n|_1^* \ge b) + \frac{18}{c}\|H_n''\|_\infty \|S_1\|_1$$

for each $c > 0$. For every chosen $b > 0$ and $c > 0$ we see that as n tends to ∞ the right hand side converges to $\frac{18b}{c}\|S_1\|_1$ which is small for appropriate choices of the constants b, c. Consequently Cauchy sequences in $\mathbb{L}$ are mapped to Cauchy sequences in the Emery topology, which – due to completeness – converge to a semi-martingale.

Lemma

Let S^n be a sequence of martingales such that $|\Delta S^n_\tau| \leq |\Delta Y_\tau|$, for all $n \geq 1$ and all stopping times τ, for some martingale Y, and let $E[[S^n, S^n] \wedge 1] \to 0$ as $n \to \infty$, then $S^n \to 0$ in the Emery topology.

Proof

Consider an arbitrary sequence $K^n \in \mathbb{S}$ of simple, predictable processes bounded by 1. We show first that $E[|(K^n \bullet S^n)_1| \wedge 1] \to 0$. We first observe that also

$$E[[(K^n \bullet S^n), (K^n \bullet S^n)]_1 \wedge 1] \to 0,$$

since K^n is uniformly bounded by 1. We use $M^n := (K^n \bullet S^n)$ as abbreviation and select a subsequence n_k such that $\mathbb{P}(|M_{n_k}| \geq 2^{-k}) \leq 2^{-k}$ for all $k \geq 1$, then $A := \sum_k [M^{n_k}, M^{n_k}]$ is almost surely finite by Borel-Cantelli.

Proof

We can consider the stopping time

$$\tau_m := \inf\{t \mid A_t \geq m \text{ or } |Y_t| \geq m\} \wedge 1\,,$$

which apparently leads to

$$[M^{n_k}_{\tau_m}, M^{n_k}_{\tau_m}] \leq A_{\tau_m} + (\Delta M^{n}_{\tau_m})^2 \leq m + (\Delta Y_{\tau_m})^2 \leq m + (m + |Y_{\tau_m}|)^2$$

for each n_k, which leads – after taking square-roots – to $E[|M^{n_k}_{\tau_m}|^*_1] \to 0$ for $k \to \infty$ by Davis inequality. Since $\mathbb{P}(\tau_m = \infty) \to 1$ as $m \to \infty$ we do also have that $(K^{n_k} \bullet S^{n_k}) \to 0$ in probability. However this already yields the result, since we have proved that every sequence $((K^n \bullet S^n))_{n \geq 0}$ has a subsequence which converges. Applying this result twice, we see that $(K^n \bullet S^n) \to 0$ in probability for all bounded by 1 simple predictable strategies, which characterizes convergence in the Emery topology.

Corollary

Let S be a martingale and $H^n \to 0$ a sequence of simple, predictable bounded by 1 strategies, then $S^n = (H^n \bullet S)$ converges to 0 in the Emery topology.

Proof.

Apparently the conditions on jumps and martingality are satisfied. It remains to show that $E[[S^n, S^n] \wedge 1] \to 0$, which is true since $[S^n, S^n] = ((H^n)^2 \bullet [S, S]) \to 0$ almost surely by the existence of $[S, S]$ and the fact that it is pathwise of finite variation. But from almost sure convergence we conclude convergence in probability. □

Preservation of martingality

Theorem

Let S be a martingale, then there is a unique continuous linear map $(. \bullet S)$ from bounded, predictable processes with respect to the uniform topology to semi-martingales $\mathbb{S}$ in the Emery topology extending J_S such that dominated convergence holds true, i.e. if $H_n - H_m \to 0$ pointwise, as $n, m \to \infty$ for a sequence of simple, predictable strategies with $\|H_n\|_\infty \leq 1$, for $n \geq 1$, then $((H_n - H_m) \bullet S) \to 0$ in the Emery topology as $n, m \to \infty$. J_S takes values in the set of local martingales.

Proof

This can be seen by the following proof: for every subsequence $m_n \geq n$ we have that $((H_n - H_{m_n}) \bullet S) \to 0$ in the Emery topology, hence $(H_n - H_m \bullet S) \to 0$ as $n, m \to \infty$.

The extension of J_S is defined by considering almost surely converging sequences $H_n \to H$ in $\mathbb{S}$ being uniformly bounded, which yield – by the previous dominated convergence result – Cauchy sequences $((H_n \bullet S))_{n \geq 1}$ in the Emery topology. This, however, means that the limits are semi-martingales, uniquely defined and linearly depending on H. Finally the resulting map is continuous.

This result immediately generalizes to semi-martingales: let S be a semi-martingale, then there is a unique continuous linear map $(. \bullet S)$ from bounded, predictable processes with respect to the uniform norm to semi-martingales $\mathbb{S}$ in the Emery topology extending J_S such that dominated convergence holds true: if $H_n \to 0$ pointwise for $\|H_n\| \leq 1$, for $n \geq 1$, then $(H_n \bullet S) \to 0$ in the Emery topology.

It is sufficient to see the statement for local martingales: we have to show that bounded convergence holds for local martingales. Let H^n be a sequence of simple, predictable bounded by 1 strategies converging almost surely to 0 and let τ_m be a localizing sequence for S, then $(K^n H^n \bullet S^{\tau_m}) \to 0$ as $n \to \infty$ in probability for all simple, predictable and bounded by 1 sequences K^n. This means that $(K^n H^n \bullet S) \to 0$ in probability, which in turn yields the statement.

Finally this leads us to the following structure: let S be a semi-martingale, then $H \mapsto (H \bullet S)$ is a continuous map, where we consider pathwise uniform convergence in probability on the set of bounded predictable strategies and the Emery topology on the set of semi-martingales. We can re-define the defining metric for the Emery topology by taking the supremum over all bounded predictable strategies. By the previous continuous extension result both metrics coincide, since every value $(K \bullet S)$ can be approximated by values $(K_n \bullet S)$ where K_n is bounded, simple and predictable.

Additionally we have the property that $L^\infty_{pred}(\Omega \times [0,1]) \times \mathbb{SEM} \to \mathbb{SEM}$, $(H,S) \mapsto (H \bullet S)$ is continuous by definition of the Emery topology.

It is our final goal, after having achieved a characterization of good integrators and a stochastic integral for bounded predictable strategies to create the somehow largest set of integrands for a given semi-martingale.

Definition

Let H be a predictable process: consider $H_n := H1_{\{\|H\| \leq n\}}$, for $n \geq 1$. If $(H_n \bullet S)$ is a Cauchy sequence in the Emery topology, then we call H integrable with respect to S, in signs $H \in L(S)$ and we write $(H \bullet S) = \lim_{n \to \infty}(H_n \bullet S)$.

Notice in particular that $(H^2 \bullet [S, S])$ is a well defined cadlag process of finite variation, which explains the connection with the previous notation.

By the very definition of the Emery topology the following lemma is clear: let $(H_n \bullet S) \to 0$ in the Emery topology and $|K_n| \leq |H_n|$, for $n \geq 0$, then also $(K_n \bullet S) \to 0$. Notice that we use here that the supremum goes over all predictable strategies, so changing signs works.

Theorem

Let S be a semi-martingale. Then $H \in L(S)$ if and only if H is predictable and for all sequences $(K_n)_{n\geq 0}$ of bounded, predictable processes with $|K_n| \leq H$ and $K_n \to 0$ pointwise, it holds that $(K_n \bullet S) \to 0$.

Proof

Let $H \in L(S)$ be fixed, then we know that $H_n := H1_{\|H\| \leq n}$, for $n \geq 1$ leads to a converging sequence $(H_n \bullet S) \to (H \bullet S)$ in the Emery topology. Take a sequence $(K_n)_{n \geq 0}$ of bounded, predictable processes with $|K_n| \leq H$ and $K_n \to 0$ pointwise, then we can find a subsequence which converges in the Emery topology.

Consider a partition of unity

$$1 = \sum_{n \geq 1} 1_{\{n-1 \leq |H| < n\}} \, .$$

For a given sequence $m_k \geq k$, $k \geq 1$ the cut-off sums

$$R_k := \sum_{k \leq n \leq m_k} H1_{\{n-1 \leq |H| < n\}} \to 0$$

as $k \to \infty$.

Proof

Furthermore by Cauchy property of $((H_n \bullet S))_{n\geq 0}$, we obtain $(R_k \bullet S) \to 0$ with respect to the Emery topology. Hence

$$\Big(\sum_{k\leq n\leq m_k} K_n 1_{\{n-1\leq |H|<n\}} \bullet S \Big) \to 0$$

as $k \to \infty$, which translates to $(K_n 1_{\{\|H\|\leq n\}} \bullet S) = (K_n \bullet (1_{\{\|H\|\leq n\}} \bullet S)) \to 0$. Since $(1_{\{\|H\|\leq n\}} \bullet S)) \to S$ in the Emery topology we arrive at the result by joint continuity of the stochastic integral.

Vice versa: assume that we have H predictable satisfying the above properties and take the previous partition of unity. Then $(R_k \bullet S) \to 0$ in the Emery topology as $k \to \infty$ for any sequence $m_k \geq k$, for $k \geq 1$. This, however, means that $((H_n \bullet S))_{n\geq 0}$ forms a Cauchy sequence.

Vector-valued stochastic integration needs some care since we do not have the usual additivity $(\sum \varphi^i \bullet S^i) = \sum(\varphi^i \bullet S^i)$ in general. A careful, clear and quick introduction is given in Cherny-Shiryaev.