Reading Group: Probability With Martingales Ch12

LEUNG Man Fung, Heman

Summer 2020

Martingales bounded in $\mathcal{L}^2$

Introduction

- Boundedness of a martingale is important for checking convergence
    - Yet boundedness in $\mathcal{L}^1$ can be difficult to check
    - Boundedness in $\mathcal{L}^1$: $\sup_n E(|M_n|) < \infty$
    - What is the difference between boundedness in $\mathcal{L}^1$ and integrability $E(|M_n|) < \infty, \forall n$?
- A martingale M bounded in $\mathcal{L}^2$ is also bounded in $\mathcal{L}^1$
    - Easier to check boundedness in $\mathcal{L}^2$ due to a Pythagorean formula

$$E(M_n^2) = E(M_0^2) + \sum_{k=1}^{n} E\left[(M_k - M_{k-1})^2\right]$$

- This chapter also presents neat proofs of:
    - Three-Series Theorem
    - Strong Law of Large Numbers
    - Lévy's extension of the Borel-Cantelli Lemmas

Martingales in $\mathcal{L}^2$: orthogonal increments

- Let $M = \{M_n\}_{n\geq 0}$ be a martingale in $\mathcal{L}^2$ so that $E(M_n^2) < \infty, \forall n$
- By martingale property, for positive integers $s \leq t \leq u \leq v$, we have

$$E(M_v|\mathcal{F}_u) = M_u \quad (a.s.)$$

- This implies the future increment $M_v - M_u$ is orthogonal to the present information $\mathcal{L}^2(\mathcal{F}_u)$, so

$$\langle M_t - M_s, M_v - M_u \rangle = 0$$

  - Future increment is also orthogonal to the past increment since $M_t - M_s \in \mathcal{L}^2(\mathcal{F}_u)$
- Hence it is possible to express M_n by sum of orthogonal increments:

$$M_n = M_0 + \sum_{k=1}^{n} (M_k - M_{k-1})$$

- Pythagoras's theorem yields (since expectation of cross term vanishes)

$$E(M_n^2) = E(M_0^2) + \sum_{k=1}^{n} E\left[(M_k - M_{k-1})^2\right]$$

Boundedness in $\mathcal{L}^2$: sum of increments square

- Theorem 12.1.1 (numbered by order in the section):
  - Let M be a martingale for which $M_n \in \mathcal{L}^2, \forall n$
  - Then M is bounded in $\mathcal{L}^2$ if and only if $\sum E\left[(M_k - M_{k-1})^2\right] < \infty$
  - And when this obtains, $M_n \to M_\infty$ almost surely and in $\mathcal{L}^2$
    - Note: William implicitly assumed the martingale was indexed in discrete time by using $k-1$
    - However I think this theorem also holds for continuous time
- Proof of $\sup_n E(M_n^2) < \infty \iff \sum E\left[(M_k - M_{k-1})^2\right] < \infty$
  - Use the Pythagorean formula

$$E(M_n^2) = E(M_0^2) + \sum_{k=1}^{n} E\left[(M_k - M_{k-1})^2\right]$$

    - Note: $E(M_0^2)$ is unbounded implies $E\left[(M_1 - M_0)^2\right]$ and $E(M_n^2)$ are also unbounded
    - So the theorem is safe even if there is no $E(M_0^2)$ explicitly

- Proof of $M_n \to M_\infty$ almost surely and in $\mathcal{L}^2$
  - Suppose that M is bounded in $\mathcal{L}^2$
  - By monotonicity of norms, M is also bounded in $\mathcal{L}^1$
  - Apply Doob's convergence theorem, we have $M_n \overset{a.s.}{\to} M_\infty$
  - The Pythagorean formula implies that $E\left[(M_{n+r} - M_n)^2\right] = \sum_{k=n+1}^{n+r} E\left[(M_k - M_{k-1})^2\right]$
  - When $r \to \infty$, Fatou's lemma yields $E\left[(M_\infty - M_n)^2\right] \leq \sum_{k \geq n+1} E\left[(M_k - M_{k-1})^2\right]$
  - Hence $\lim_n E\left[(M_\infty - M_n)^2\right] = 0$, i.e. $M_n \overset{\mathcal{L}^2}{\to} M_\infty$
    - Intuition: when $n \to \infty$, there is no more increment on RHS

Sum of independent random variables in $\mathcal{L}^2$

Sum of independent zero-mean RVs in $\mathcal{L}^2$

- Theorem 12.2.1:
  - Suppose that $\{X_k\}_{k\in\mathbb{N}}$ is a sequence of independent RVs with zero-mean and finite variance σ_k^2
  - Then $\sum \sigma_k^2 < \infty \implies \sum X_k$ converges almost surely
  - Further if X_k is bounded by some positive constant K, then the reverse direction is also true
    - i.e. $\sum X_k$ converges almost surely $\implies \sum \sigma_k^2 < \infty$
- Notation: define
  - Natural filtration: $\mathcal{F}_n := \sigma(X_1, X_2, \ldots, X_n)$ where $\mathcal{F}_0 := \{\varnothing, \Omega\}$
  - Partial sum: $M_n := \sum_{k=1}^n X_k$ where $M_0 := 0$
  - $A_n := \sum_{k=1}^n \sigma_k^2$ where $A_0 := 0$
  - $N_n := M_n^2 - A_n$ where $N_0 := 0$

- Proof of $\sum \sigma_k^2 < \infty \implies \sum X_k$ converges almost surely
  - From example in 10.4, M is a martingale
  - Using the Pythagorean formula,

$$E(M_n^2) = \sum_{k=1}^{n} E\left[(M_k - M_{k-1})^2\right] = \sum_{k=1}^{n} E(X_k^2) = \sum_{k=1}^{n} \sigma_k^2 = A_n$$

  - If $\sum \sigma_k^2 < \infty$, then M is bounded in $\mathcal{L}^2$ and M_n converges almost surely by theorem 12.1.1

- Proof of $\sum X_k$ converges almost surely $\implies \sum \sigma_k^2 < \infty$
  - Since $X_k \perp \mathcal{F}_{k-1}$, we have, almost surely

$$E\left[(M_k - M_{k-1})^2 | \mathcal{F}_{k-1}\right] = E[X_k^2 | \mathcal{F}_{k-1}] = E(X_k^2) = \sigma_k^2$$

  - Similarly, since M_{k-1} is $\mathcal{F}_{k-1}$ measurable, we can expand $(M_k - M_{k-1})^2$, almost surely

$$\sigma_k^2 = E(M_k^2 | \mathcal{F}_{k-1}) - 2M_{k-1}E(M_k | \mathcal{F}_{k-1}) + M_{k-1}^2 = E(M_k^2 | \mathcal{F}_{k-1}) - M_{k-1}^2$$

  - But this implies that N is a martingale (Recall $N_n := M_n^2 - A_n$)
  - Now let $c \in (0, \infty)$ and $T := \inf\{r : |M_r| > c\}$
  - Since stopped martingale is also a martingale, $E(N_n^T) = E\left[(M_n^T)^2\right] - E(A_{T \wedge n}) = 0$
  - By the further condition, we have $|M_T - M_{T-1}| = |X_T| \leq K$ if $T < \infty$
  - Hence $E(A_{T \wedge n}) = E\left[(M_n^T)^2\right] \leq (K + c)^2, \forall n$
    - Intuition: same as upcrossing with last increment bounded by K
  - However, since $\sum X_k$ converges a.s., the partial sums are a.s. bounded
  - So it must be the case that $P(T = \infty) > 0$ for some c and $A_\infty := \sum \sigma_k^2 < \infty$

Random signs

- Let $\{a_n\}$ be a sequence of real numbers and $\{\epsilon_n\}$ be a sequence of iid Rademacher RVs
  - Rademacher distribution: $P(\epsilon_n = \pm 1) = 0.5$
  - Frequently appear in statistical learning theory
- Theorem 12.2.1 tells us that $\sum \epsilon_n a_n$ converges a.s. $\iff \sum a_n^2 < \infty$
  - And $\sum \epsilon_n a_n$ oscillates infinitely if $\sum a_n^2 = \infty$
- Sketch
  - Note that $Var(\epsilon_k a_k) = a_k^2$ and $|\epsilon_k a_k| \leq \sup_n a_n$, theorem 12.2.1 will yield the first part
    - $\sup_n a_n < \infty$ because we are given $\sum a_n^2 = \infty$
  - For the second part, my guess is since $\sum a_n^2 = \infty$, $\sum \epsilon_n a_n$ will not converge
  - However, as ϵ_n are Rademacher RVs, $\sum \epsilon_n a_n$ will oscillate depending on the realization

Symmetrization: expanding the sample space

- What if the mean of RVs is non-zero?
- Lemma 12.4.1
  - Suppose $\{X_n\}$ is a sequence of independent RVs bounded by a constant $K \in [0, \infty)$
  - Then $\sum X_n$ converges a.s. implies that $\sum E(X_n)$ converges and $\sum Var(X_n) < \infty$
- Proof
  - If $E(X_n) = 0, \forall n$, then this reduce to theorem 12.2.1
  - Otherwise we need to replace each X_n by a "symmetrized version" Z_n^* of mean 0
  - Let $\left(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}, (\tilde{X}_n : n \in \mathbb{N})\right)$ be an exact copy of $\left(\Omega, \mathcal{F}, \mathbb{P}, (X_n : n \in \mathbb{N})\right)$
  - Define a richer probability space $\left(\Omega^*, \mathcal{F}^*, \mathbb{P}^*\right) := \left(\Omega, \mathcal{F}, \mathbb{P}\right) \times \left(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}\right)$
  - For $\omega^* = (\omega, \tilde{\omega}) \in \Omega$, define

$$X_n^*(\omega^*) := X_n(\omega), \tilde{X}_n^*(\omega^*) := \tilde{X}_n(\tilde{\omega}), Z_n^*(\omega^*) := X_n^*(\omega^*) - \tilde{X}_n^*(\omega^*)$$

    - Intuition: X_n^* is X_n lifted to the richer probability space

- Proof (continue)
  - It is clear that the combined family $(X_n : n \in \mathbb{N}) \cup (\tilde{X}_n : n \in \mathbb{N})$ is on $(\Omega^*, \mathcal{F}^*, \mathbb{P}^*)$
    - This may be proved by the uniqueness lemma in 1.6
  - Both $X_n^*, \tilde{X}_n^*$ having the same $\mathbb{P}^*$-distribution as the $\mathbb{P}$-distribution of X_n

$$\mathbb{P}^* \circ (X_n^*)^{-1} = \mathbb{P} \circ X_n^{-1} \text{ on } (\mathbb{R}, \mathcal{B}), \text{etc.}$$

  - Now $(Z_n^* : n \in \mathbb{N}^*)$ is a zero-mean sequence of independent RVs on $(\Omega^*, \mathcal{F}^*, \mathbb{P}^*)$
  - We have $|Z_n^*(\omega^*)| \le 2K, \forall n, \forall \omega^*$ and $Var(Z_n^*) = 2\sigma_n^2$ where $\sigma_n^2 := Var(X_n)$
    - This is probably due to independence of original RV and its copy
  - Let $G := \{\omega \in \Omega : \sum X_n(\omega) \text{ converges}\}$ with $\tilde{G}$ defined similarly
  - Since $\mathbb{P}(G) = \tilde{\mathbb{P}}(\tilde{G}) = 1$, $\mathbb{P}^*(G \times \tilde{G}) = 1$
  - But $\sum Z_n^*(\omega^*)$ also converges on $G \times \tilde{G}$, which means $\mathbb{P}^*(\sum Z_n^* \text{ converges}) = 1$
  - As Z_n^* converges a.s., is zero-mean and bounded, theorem 12.2.1 yields $\sum \sigma_n^2 < \infty$
  - It also follows that $\sum[X_n - E(X_n)]$ and $\sum E(X_n)$ converges a.s.

Some lemmas on real numbers

Cesàro's lemma

- Alternative version of Stolz–Cesàro theorem
- Suppose that $\{b_n\}$ is a sequence of strictly positive real numbers with $b_0 := 0$ and $b_n \uparrow \infty$
- $\{v_n\}$ is a convergent sequence of real numbers with $v_n \to v_\infty \in \mathbb{R}$
- Then we have $\lim_{n\to\infty} \frac{1}{b_n} \sum_{k=1}^{n} (b_k - b_{k-1}) v_k = v_\infty$
- Proof: let $\epsilon > 0$. Choose N s.t. $v_k > v_\infty - \epsilon$ whenever $k \geq N$. Then

$$\begin{aligned}\liminf_{n\to\infty} \frac{1}{b_n} \sum_{k=1}^{n} (b_k - b_{k-1}) v_k &\geq \liminf_{n\to\infty} \left[\frac{1}{b_n} \sum_{k=1}^{N} (b_k - b_{k-1}) v_k + \frac{b_n - b_N}{b_n} (v_\infty - \epsilon) \right] \\ &\geq 0 + v_\infty - \epsilon\end{aligned}$$

- Since this is true for every $\epsilon > 0$, we have $\liminf \geq v_\infty$
- By a similar argument, we have $\limsup \leq v_\infty$ and the result follows

Kronecker's lemma

- Suppose that $\{b_n\}$ is a sequence of strictly positive real numbers with $b_n \uparrow \infty$
- $\{x_n\}$ is a sequence of real numbers and define $s_n := \sum_{i=1}^{n} x_i$
- Then we have $\sum \frac{x_n}{b_n}$ converges $\implies \frac{s_n}{b_n} \to 0$
- Proof: let $u_n := \sum_{k \le n} \frac{x_k}{b_k}$ so that $u_\infty := \lim_{n\to\infty} u_n$ exists
- Then $u_n - u_{n-1} = \frac{x_n}{b_n}$. Thus by rearrangement

$$s_n = \sum_{k=1}^{n} b_k(u_k - u_{k-1}) = b_n u_n - \sum_{k=1}^{n} (b_k - b_{k-1}) u_{k-1}$$

- Applying Cesàro's lemma, we have $\frac{s_n}{b_n} \to u_\infty - u_\infty = 0$
- Alternative version: $\sum x_n$ exists and is finite $\implies \lim_{n\to\infty} \frac{1}{b_n} \sum_{k=1}^{n} b_k x_k = 0$
  - Check the little o of a weighted sum with monotonically increasing weights

Some neat proofs of classical theorems

Kolmogorov's Three-Series Theorem

- Let $\{X_n\}$ be a sequence of independent RVs
- Then $\sum X_n$ converges a.s. iff for some (then for every) $K > 0$, the following 3 properties hold:
  - $\sum_n P(|X_n| > K) < \infty$
  - $\sum_n E(X_n^K)$ converges
  - $\sum_n Var(X_n^K) < \infty$ where

$$X_n^K(\omega) := \begin{cases} X_n(\omega) & , |X_n(\omega)| \leq K \\ 0 & , |X_n(\omega)| > K \end{cases}$$

- Proof of "only if" part
  - Suppose that $\sum X_n$ converges a.s. and K is any constant in $(0, \infty)$
  - Since $X_n \to 0$ a.s. whence $|X_n| > K$ for only finitely many n, BC2 shows the first property holds
    - BC2: $\sum P(|X_n| > K) = \infty \implies P(|X_n| > K, \text{ i.o.}) = 1$
    - Contraposition: $P(|X_n| > K, \text{ i.o.}) = 0 \implies \sum P(|X_n| > K) < \infty$
  - Since (a.s.) $X_n = X_n^K$ for all but finitely many n, $\sum X_n^K$ also converges a.s.
  - Applying lemma 12.4.1 yields the other two properties

- Proof of “if” part
  - Suppose that for some $K > 0$ the 3 properties hold
  - Then $\sum P(X_n \neq X_n^K) = \sum P(|X_n| > K) < \infty$ by construction and property 1
  - Applying BC1 yields $P(X_n = X_n^K \text{ for all but finitely many } n) = 1$
  - So we only need to check $\sum X_n^K$ converges a.s.
  - By property 2, we can check if $\sum \left[X_n^K - E(X_n^K)\right]$ converges a.s. instead
  - Now note that $Y_n^K := X_n^K - E(X_n^K)$ is a zero-mean RV with $E\left[(Y_n^K)^2\right] = Var(X_n^K)$
  - By property 3, the result follows from theorem 12.2.1

A Strong Law under variance constraints

- Lemma 12.8.1
  - Let $\{W_n\}$ be a sequence of independent RVs with $E(W_n) = 0, \sum \frac{Var(W_n)}{n^2} < \infty$
  - Then $\frac{1}{n} \sum_{k \leq n} W_k \overset{a.s.}{\to} 0$
- Proof
  - By Kronecker's lemma, it suffices to prove that $\sum \frac{W_n}{n}$ converges
  - However $E\left(\frac{W_n}{n}\right) = 0, \sum Var\left(\frac{W_n}{n}\right) = \sum \frac{Var(W_n)}{n^2} < \infty$
  - So by theorem 12.2.1, the statement is proved

Kolmogorov's Truncation Lemma

- Suppose that $X_1, X_2, ...$ are iid RVs with the same distribution as X where $E(|X|) < \infty$
- Define

$$\mu := E(X), Y_n := \begin{cases} X_n & , |X_n| \leq n \\ 0 & , |X_n| > n \end{cases}$$

- Then
  - $E(Y_n) \to \mu$
  - $P(Y_n = X_n \text{ eventually}) = 1$
  - $\sum \frac{Var(Y_n)}{n^2} < \infty$

- Proof of $E(Y_n) \to \mu$
  - Let

$$Z_n := \begin{cases} X & , |X| \le n \\ 0 & , |X| > n \end{cases}$$

  - Then $Z_n \overset{d}{=} Y_n$ and $E(Z_n) = E(Y_n)$
  - When $n \to \infty$, we have $Z_n \to X, |Z_n| \le |X|$
  - Applying dominated convergence theorem (note that X is integrable by assumption):

$$\lim_{n\to\infty} E(Y_n) = \lim_{n\to\infty} E(Z_n) = E(X) = \mu$$

- Proof of $P(Y_n = X_n \text{ eventually}) = 1$
  - Note that

$$\begin{aligned}\sum_{n=1}^{\infty} P(Y_n \neq X_n) = \sum_{n=1}^{\infty} P(|X_n| > n) = \sum_{n=1}^{\infty} P(|X| > n) \\ = E\left(\sum_{n=1}^{\infty} I_{|X|>n}\right) = E\left(\sum_{1 \leq n < |X|} 1\right) \\ \leq E(|X|) < \infty\end{aligned}$$

  - By BC1, $P(Y_n \neq X_n, \text{ i.o}) = 0$. In other words, $P(Y_n = X_n, \text{ e.v.}) = 1$

- Proof of $\sum \frac{Var(Y_n)}{n^2} < \infty$
  - We have

$$\sum \frac{Var(Y_n)}{n^2} \leq \sum \frac{E(Y_n^2)}{n^2} = \sum_n \frac{E(|X|^2; |X| \leq n)}{n^2} = E\left[|X|^2 f(|X|)\right]$$

    - where $f(z) = \sum_{n \geq \max(1,z)} \frac{1}{n^2}, 0 < z < \infty$
  - Note that, for $n \geq 1$, $\frac{1}{n^2} \leq \frac{2}{n(n+1)} = 2\left(\frac{1}{n} - \frac{1}{n+1}\right)$
  - Hence $f(z) \leq \frac{2}{\max(1,z)}$ by telescoping
  - We have $\sum \frac{Var(Y_n)}{n^2} \leq 2E(|X|) < \infty$

Kolmogorov's Strong Law of Large Numbers

- Let $X_1, X_2, \dots$ be iid RVs with $E(|X_k|) < \infty, \forall k$. Define $S_n := \sum_{k=1}^{n} X_k$ and $\mu := E(X_k), \forall k$
- Then $\frac{1}{n} S_n \overset{a.s.}{\to} \mu$
- Proof
  - Define Y_n as in Kolmogorov's Truncation Lemma
  - By $P(Y_n = X_n, \text{ e.v.}) = 1$, it suffices to show that $\frac{1}{n} \sum_{k=1}^{n} Y_k \overset{a.s.}{\to} \mu$
  - Define $W_k := Y_k - E(Y_k)$. Note that

$$\frac{1}{n} \sum_{k=1}^{n} Y_k = \frac{1}{n} \sum_{k=1}^{n} E(Y_k) + \frac{1}{n} \sum_{k=1}^{n} W_k$$

  - The first term $\frac{1}{n} \sum_{k=1}^{n} E(Y_k) \to \mu$ by $E(Y_n) \to \mu$ and Cesàro's lemma (let $b_n := n$)
  - The second term $\frac{1}{n} \sum_{k=1}^{n} W_k \overset{a.s.}{\to} 0$ by $\sum \frac{Var(Y_n)}{n^2} < \infty$ and lemma 12.8.1

Some remarks on SLLN

- Philosophy
  - SLLN gives a precise formulation of $E(X)$ as "the mean of a large number of independent realizations of X"
    - Long run guarantee of frequentist method
  - From exercise E4.6, it can be shown that if $E(|X|) = \infty$, then $\limsup \frac{S_n}{n} = \infty$ almost surely
  - Hence SLLN is the best possible result for iid RVs
- Methodology
  - The truncation technique seems "ad hoc" with no pure-mathematical elegance
  - The proof with martingale or ergodic theory possess that
  - However, each of the methods can be adapted to cover situations which the others cannot tackle
  - Classical truncation arguments retain great importance

Decomposition of stochastic process

Doob decomposition

- Theorem 12.11.1
  - Let $\{X_n\}_{n\in\mathbb{Z}^+}$ be an adapted process in $\mathcal{L}^1$
  - Then X has a Doob decomposition $X = X_0 + M + A$
    - where M is a martingale null at 0 and A is a previsible process null at 0
  - Moreover, this decomposition is unique modulo indistinguishability in the sense that

$$X = X_0 + \tilde{M} + \tilde{A} \implies P(M_n = \tilde{M}_n, A_n = \tilde{A}_n, \forall n) = 1$$

  - Continuous time analogue: Doob-Meyer decomposition
- Corollary 12.11.2
  - X is a submartingale iff A is an increasing process in the sense that $P(A_n \leq A_{n+1}, \forall n) = 1$
  - Similarly, X is a supermartingale if and only if A is almost surely decreasing

- Proof of existence
  - If X has Doob decomposition $X = X_0 + M + A$, we have

$$\begin{aligned} E(X_n - X_{n-1}|\mathcal{F}_{n-1}) &= E(M_n - M_{n-1}|\mathcal{F}_{n-1}) + E(A_n - A_{n-1}|\mathcal{F}_{n-1}) \\ &= 0 + (A_n - A_{n-1}) \end{aligned}$$

  - Hence we can define A by $A_n = \sum_{k=1}^{n} E(X_k - X_{k-1}|\mathcal{F}_{n-1})$ a.s.
    - A represents the sum of expected increments of X
    - M can be defined by $M_n = \sum_{k=1}^{n} [X_k - E(X_k|\mathcal{F}_{k-1})]$, which adds up the surprises
  - Corollary is now obvious by the defintion of A
- Proof of uniqueness
  - Define $Y := M - \tilde{M} = A - \tilde{A}$ by rearranging the other decomposition
  - The first equality implies that Y is a martingale and $E(Y_n|\mathcal{F}_{n-1}) = Y_{n-1}$ a.s.
  - The second equality implies that Y is also previsible and $E(Y_n|\mathcal{F}_{n-1}) = Y_n$ a.s.
  - Since $Y_0 = 0$ by construction, this implies that $Y_n = 0$ a.s.
  - which also means that the decomposition is almost surely unique

The angle-brackets process $\langle M \rangle$

- Let M be a martingale in $\mathcal{L}^2$ and null at 0
- The the conditional form of Jensen's inequality shows that M^2 is a submartingale
  - Square function is convex as the second derivative is non-negative
  - $E(M_n^2|\mathcal{F}_{n-1}) \geq [E(M_n|\mathcal{F}_{n-1})]^2 = M_{n-1}^2$
- Thus M^2 has a Doob decomposition $M^2 = N + A$
  - where N is a martingale null at 0 and A is a previsible increasing process null at 0
  - A is often written as $\langle M \rangle$ (quadratic variation in stochastic calculus)
- Since $E(M_n^2) = E(A_n)$, M is bounded in $\mathcal{L}^2 \iff E(A_\infty) < \infty$
  - where $A_\infty := \uparrow \lim A_n$, a.s.
  - $E(N) = E\left[E(N|\mathcal{F}_0)\right] = 0$ (martingale property)
- It is important to note that $A_n - A_{n-1} = E(M_n^2 - M_{n-1}^2|\mathcal{F}_{n-1}) = E\left[(M_n - M_{n-1})^2|\mathcal{F}_{n-1}\right]$
  - As the cross term is $-E(2M_nM_{n-1}|\mathcal{F}_{n-1}) = -2M_{n-1}^2$

Relating convergence of M to finiteness of $\langle M \rangle_\infty$

- Theorem 12.13.1
  - Let M be a martingale in $\mathcal{L}^2$ and null at 0. Let A be "a version of" $\langle M \rangle$
  - Then $A_\infty(\omega) < \infty \implies \lim_{n\to\infty} M_n(\omega)$ exists
  - Suppose that M has uniformly bounded increments in that for some $K \in \mathbb{R}$,

$$|M_n(\omega) - M_{n-1}(\omega)| \leq K, \forall n, \forall \omega$$

  - Then $\lim_{n\to\infty} M_n(\omega)$ exists $\implies A_\infty(\omega) < \infty$
- Remark
  - Theorem 12.13.1 is an extension of 12.2.1
    - Doob convergence theorem + 12.2.1 with different conditions

- Proof of $A_\infty(\omega) < \infty \implies \lim_{n\to\infty} M_n(\omega)$ exists
  - Since A is previsible, $S(k) := \inf\left\{n \in \mathbb{Z}^+ : A_{n+1} > k\right\}$ is a stopping time for every $k \in \mathbb{N}$
  - The stopped process $A^{S(k)}$ is also previsible because for $B \in \mathcal{B}, n \in \mathbb{N}$

$$\left\{A_{n\wedge S(k)} \in B\right\} = F_1 \cup F_2$$

    - where $F_1 := \cup_{r=0}^{n-1}\left\{S(k) = r; A_r \in B\right\} \in \mathcal{F}_{n-1}$ (case $S(k) \leq n$)
    - and $F_2 := \left\{A_n \in B\right\} \cap \left\{S(k) \leq n-1\right\}^c \in \mathcal{F}_{n-1}$ (case $S(k) > n$)
  - Since $\left(M^{S(k)}\right)^2 - A^{S(k)} = (M^2 - A)^{S(k)}$ is a martingale, we have $\langle M^{S(k)}\rangle = A^{S(k)}$
    - Why this is not true by definition?
  - As $A^{S(k)}$ is bounded by k, $M^{S(k)}$ is bounded in $\mathcal{L}^2$ by the third property in 12.2
  - Thus $\lim_n M_{n\wedge S(k)}$ exists almost surely by Doob convergence theorem
  - However, $\left\{A_\infty < \infty\right\} = \cup_k\left\{S(k) = \infty\right\}$
  - The result now follows on combining $\lim_n M_{n\wedge S(k)}$ and $\left\{A_\infty < \infty\right\}$

- Proof of $\lim_{n\to\infty} M_n(\omega)$ exists $\implies A_\infty(\omega) < \infty$
  - Suppose that $P(A_\infty = \infty, \sup_n |M_n| < \infty) > 0$
  - Then for some $c > 0$, $P[T(c) = \infty, A_\infty = \infty] > 0$ (since M_n is bounded)
    - where $T(c) := \inf \{r : |M_r| > c\}$ is a stopping time
  - Now $E\left[M^2_{T(c)\wedge n} - A_{T(c)\wedge n}\right] = 0$ and $M^{T(c)}$ is bounded by $c + K$
    - The first one comes from decomposition and martingale property
    - The second one comes from the given condition and idea of upcrossing
  - Thus $E\left[A_{T(c)\wedge n}\right] \le (c+K)^2, \forall n$, which implies $E(A_\infty) < \infty$
  - Contradication arises so we should have $P(A_\infty = \infty, \sup_n |M_n| < \infty) = 0$
- Remarks
  - The additional assumption of uniformly bounded increments of M is needed for upcrossing
  - For A, this is not necessary as the jump $A_{S(k)} - A_{S(k)-1}$ becomes irrelevant due to previsibility

A trivial “Strong Law” for martingales in $\mathcal{L}^2$

- Let M be a martingale in $\mathcal{L}^2$ and null at 0. Let A be “a version of” $\langle M \rangle$
- Since $(1+A)^{-1}$ is a bounded previsible process, we can define a martingale

$$W_n := \sum_{k=1}^{n} \frac{M_k - M_{k-1}}{1 + A_k} = \left[(1+A)^{-1} \bullet M\right]_n$$

- Moreover, since $(1 + A_n)$ is $\mathcal{F}_{n-1}$ measurable,

$$\begin{aligned} E\left[(W_n - W_{n-1})^2 | \mathcal{F}_{n-1}\right] &= (1 + A_n)^{-2}(A_n - A_{n-1}) \\ &\leq (1 + A_{n-1})^{-1} - (1 + A_n)^{-1}, \text{ a.s.} \end{aligned}$$

- We see that $\langle W \rangle_\infty \leq 1$ so $\lim W_n$ exists a.s. by theorem 12.13.1
- Applying Kronecker’s lemma shows that $\frac{M_n}{A_n} \to 0$ almost surely on $\{A_\infty = \infty\}$

Lévy's extension of the Borel-Cantelli Lemmas

- Theorem 12.15.1
  - Suppose that for $n \in \mathbb{N}, E_n \in \mathcal{F}_n$
  - Define $Z_n := \sum_{k=1}^{n} I_{E_k}$ = number of $E_k (k \leq n)$ which occur
  - Also define $\xi_k := P(E_k | \mathcal{F}_{k-1})$ and $Y_n := \sum_{k=1}^{n} \xi_k$
  - Then we have $\{Y_\infty < \infty\} \implies \{Z_\infty < \infty\}$ almost surely
  - And $\{Y_\infty = \infty\} \implies \{\frac{Z_n}{Y_n} \to 1\}$ almost surely
- Extension of BC1
  - Since $E(\xi_k) = P(E_k)$, it follows that if $\sum P(E_k) < \infty$ then $Y_\infty < \infty$ a.s. and BC1 follows
- Extension of BC2
  - Let $\{E_n\}_{n \in \mathbb{N}}$ be a sequence of independent events associated with some triple $(\Omega, \mathcal{F}, \mathbb{P})$
  - Define the natural filtration $\mathcal{F}_n = \sigma(E_1, E_2, \ldots, E_n)$
  - Then $\xi_k = P(E_k)$ almost surely by independence
  - BC2 follows from $\{Y_\infty = \infty\} \implies \{\frac{Z_n}{Y_n} \to 1\}$ a.s.

- Proof
  - Let M be the martingale $Z - Y$, so that $Z = M + Y$ is the Doob decomposition of Z. Then

$$M_n = Z_n - Y_n = \sum_{k=1}^{n} [I_{E_k} - \xi_k]$$

$$A_n := \langle M \rangle_n = \sum_{k=1}^{n} E\left[(M_k - M_{k-1})^2 | \mathcal{F}_{k-1}\right] = \sum_{k=1}^{n} E\left[(I_{E_k} - \xi_k)^2 | \mathcal{F}_{k-1}\right]$$

$$= \sum_{k=1}^{n} E\left[I_{E_k} - 2I_{E_k}\xi_k + \xi_k^2 | \mathcal{F}_{k-1}\right] = \sum_{k=1}^{n} \xi_k(1 - \xi_k) \leq Y_n, \text{ a.s.}$$

    - Note that $E(I_{E_k}|\mathcal{F}_{k-1}) = P(E_k|\mathcal{F}_{k-1}) =: \xi_k$
  - If $Y_\infty < \infty$, then $A_\infty < \infty$ and $\lim M_n$ exists so that Z_∞ is finite almost surely
  - If $Y_\infty = \infty$ and $A_\infty < \infty$, then $\lim M_n$ still exists and $\frac{Z_n}{Y_n} \to 1$ almost surely
  - If $Y_\infty = \infty$ and $A_\infty = \infty$, then $\frac{M_n}{A_n} = \frac{M_n}{M_n^2 + N} \to 0$ almost surely
  - Hence, a fortiori, $\frac{M_n}{Y_n} \to 0$ and $\frac{Z_n}{Y_n} = \frac{M_n + Y_n}{Y_n} \to 1$ almost surely
    - A fortiori means “from the stronger argument”

Concluding remarks

Comments

- Independence is important in the study of RVs
- Martingale may relax the independent RVs assumption to orthogonal increments
  - Pythagorean formula in $\mathcal{L}^2$
  - Richer probability space for copy of independent RVs
  - Doob decomposition for expected increment and surprise
- Martingale also relates convergence with finiteness
  - Doob convergence theorem
  - Truncation technique with stopping time
  - $\langle M \rangle$ from decomposition of M^2
- Martingale transform is a possible candidate for control variate in variance reduction
  - Suppose $\{X_n\}_{n\in\mathbb{N}}$ is a martingale wrt natural filtration $\mathcal{F}_n$
  - $Y_{n+1} := \sum_{i=1}^{n} g_i(X_1, \ldots, X_i)(X_{i+1} - X_i)$ is also a martingale wrt $\mathcal{F}_n$
  - Choose g with high correlation to use Y_n as control variate
  - See a trivial "Strong Law" for an example of martingale transform