Lecture No 6 March 25, 2022 (Fri)

Part II. Foundations of Stochastic Analysis

§9 Discrete time martingale

9.1 Definition

- $(\Omega, \mathcal{F}, P)$: Probability space
- $(\mathcal{F}_n)_{n=1,2,\dots}$: filtration or reference family
  i.e. each $\mathcal{F}_n$ is a sub σ–field of $\mathcal{F}$ and increasing:
  $$\mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots$$
- $X = (X_n)_{n=1,2,\dots}$: discrete time real-valued stochastic process, i.e., a sequence of real-valued r.v.'s

[Definition 9.1] We call X a martingale with respect to a filtration $(\mathcal{F}_n)$ (or $(\mathcal{F}_n)$-martingale), if

(1) $(\mathcal{F}_n)$–adapted: X_n is $\mathcal{F}_n$–measurable for $^\forall n$.

(2) Integrability: $E[|X_n|] < \infty$ for $^\forall n$.

(3) For $^\forall n$, $E[X_{n+1}|\mathcal{F}_n] = X_n$ a.s. holds.

X is called submartingale if $E[X_{n+1}|\mathcal{F}_n] \geq X_n$ a.s. holds in (3) and supermartingale if $E[X_{n+1}|\mathcal{F}_n] \leq X_n$ a.s. in (3). □


Doob (from Wikipedia)

[Remark] • For $\mathcal{G}$: sub σ-field of $\mathcal{F}$, X: $E[|X|] < \infty$, recall

$$Y = E[X|\mathcal{G}] \iff \text{① } Y\text{: } \mathcal{G}\text{–measurable and}$$
$$\text{② } E[X, A] = E[Y, A],\ ^\forall A \in \mathcal{G}$$

- Therefore, (3)$\iff E[X_{n+1}, A] = E[X_n, A]$, $^\forall A \in \mathcal{F}_n$
- $m > n \geq 1 \implies E[X_m|\mathcal{F}_n] = X_n$ a.s.

∵ Since $\mathcal{F}_n \subset \mathcal{F}_{n+1}$, by Tower property, we have a.s.

$$E[X_{n+2}|\mathcal{F}_n] = E\big[E[X_{n+2}|\mathcal{F}_{n+1}]|\mathcal{F}_n\big] = E[X_{n+1}|\mathcal{F}_n] = X_n \qquad \square$$

- Expectation $E[X_n]$ is constant in n.
- X: submartingale$\iff -X$: supermartingale $\square$
  - ▶ Below, when filtration $(\mathcal{F}_n)$ is clear, it is omitted (we simply call martingale, submartingale, ...).
  - ▶ For example, for a given stochastic process X, if we take $\mathcal{F}_n = \sigma\{X_1, \ldots, X_n\}$, it is a filtration and satisfies Definition 9.1-(1): $(\mathcal{F}_n)$–adaptedness.
  - ▶ We often omit writing a.s.

[Example] (1) (Origin of martingale theory) Let $(Z_n)_{n=1,2,\dots}$ be a sequence of integrable and independent r.v.'s with mean 0: $E[X_n] = 0$. Set $\mathcal{F}_n = \sigma(Z_1, Z_2, \dots, Z_n)$ and $X_n = \sum_{k=1}^{n} Z_k$.
Then, $X = (X_n)_{n=1,2,\dots}$ is an $(\mathcal{F}_n)$–martingale.

$\because$ $E[X_{n+1} - X_n | \mathcal{F}_n] = E[Z_{n+1} | \mathcal{F}_n] \underset{\perp\!\!\!\perp}{=} E[Z_{n+1}] = 0$ □

(2) Let $(\mathcal{F}_n)$ and an integrable r.v. X be given. Then $(X_n := E[X|\mathcal{F}_n])$ is a martingale. $\because$ By Tower property. □

(3) ψ: convex function, (X_n): martingale s.t. $E[|\psi(X_n)|] < \infty$. Then, $(\psi(X_n))$ is a submartingale.

$\because$ By Jensen's inequality,

$$E\left[\psi(X_{n+1}) | \mathcal{F}_n\right] \geq \psi\left(E[X_{n+1} | \mathcal{F}_n]\right) = \psi(X_n)$$ □

In particular, for $p \geq 1$, if $E[|X_n|^p] < \infty$, then $(|X_n|^p)_{n=1,2,\dots}$ is a submartingale. □

9.2 Doob decomposition

The next theorem shows that, to study submartingales, it is enough to study martingales.

[Theorem 9.1] (Doob decomposition) A submartingale $X = (X_n)$ is decomposed into a sum of martingale part $M = (M_n)$ and increasing part $A = (A_n)$. This decomposition is unique in a.s.-sense.

(1) $X_n = M_n + A_n$

(2) (M_n) is a martingale.

(3) For $n = 2, 3, \ldots$, A_n is $\mathcal{F}_{n-1}$–measurable (called predictable) and $0 = A_1 \leq A_2 \leq \cdots a.s.$ □

P: If a submartingale X satisfies $E[X_n] = E[X_1]$ for $^\forall n$, X is a martingale.

[Proof] (Existence of decomposition) Set $M_1 = X_1$ and

$$M_n = X_1 + \sum_{k=2}^{n} \{X_k - E[X_k|\mathcal{F}_{k-1}]\}, \ n \geq 2.$$

Then, (M_n) is a martingale.

$\because$ It is clear that M_n is $\mathcal{F}_n$–measurable and integrable. Also,

$$E[M_{n+1} - M_n|\mathcal{F}_n] = E\left[X_{n+1} - E[X_{n+1}|\mathcal{F}_n]\middle|\mathcal{F}_n\right] = 0. \qquad \square$$

Next, set $A_n = X_n - M_n$. Then, obviously $A_1 = 0$. For $n = 2, 3, \ldots,$

$$\begin{aligned} A_n &= X_n - \left(X_n - E[X_n|\mathcal{F}_{n-1}] + M_{n-1}\right) \\ &= E[X_n|\mathcal{F}_{n-1}] - M_{n-1}, \end{aligned}$$

In particular, we see that A_n is $\mathcal{F}_{n-1}$–measurable. Moreover,

$$\begin{aligned} A_{n+1} - A_n &= (X_{n+1} - X_n) - (M_{n+1} - M_n) \\ &= (X_{n+1} - X_n) - (X_{n+1} - E[X_{n+1}|\mathcal{F}_n]) \\ &= E[X_{n+1}|\mathcal{F}_n] - X_n \geq 0 \text{ a.s.} \quad (\because X : \text{submartingale}) \end{aligned}$$

(Uniqueness of decomposition in a.s.-sense) If (X_n) has two decompositions $X_n = M_n + A_n = M'_n + A'_n$, we have $A_n - A'_n = M'_n - M_n$. However, since $A_n - A'_n$ is $\mathcal{F}_{n-1}$–measurable,

$$\begin{aligned} A_n - A'_n &= E[A_n - A'_n | \mathcal{F}_{n-1}] \\ &= E[M'_n - M_n | \mathcal{F}_{n-1}] \quad a.s. \end{aligned}$$

Recalling that $(M' - M)$ is a martingale,

$$(\text{RHS}) = M'_{n-1} - M_{n-1} = A_{n-1} - A'_{n-1}.$$

Repeating this, we obtain

$$A_n - A'_n = \cdots = A_1 - A'_1 = 0,$$

which implies $A_n = A'_n$ (a.s.). Thus $M_n = M'_n$ (a.s.) also follows and the uniqueness of Doob decomposition in a.s.-sense is shown. □

9.3 Markov time

We introduce a notion of random times well suited to the filtration $(\mathcal{F}_n)$.

[Definition 9.2] $\mathbb{N} \cup \{\infty\}$–valued r.v. $\tau = \tau(\omega)$ is called $(\mathcal{F}_n)$–Markov time (or stopping time)
$\underset{\text{def}}{\iff} \{\tau \le n\} \equiv \{\omega \in \Omega\,;\, \tau(\omega) \le n\} \in \mathcal{F}_n,\ {}^\forall n \in \mathbb{N}$ □

[Remark] The above condition $\iff \{\tau = n\} \in \mathcal{F}_n,\ {}^\forall n \in \mathbb{N}$

∵ $(\Longrightarrow)$ follows from $\{\tau = n\} = \{\tau \le n\} \setminus \{\tau \le n-1\}$.
$(\Longleftarrow)$ follows from $\{\tau \le n\} = \bigcup_{k=1}^{n} \{\tau = k\}$. □

[Example] (1) $\tau(\omega) \equiv m$ (constant) is a Markov time
(2) ((first) hitting time) Assume $X = (X_n)$ is $(\mathcal{F}_n)$–adapted. For $A \in \mathcal{B}(\mathbb{R})$, set

$$\tau_A(\omega) = \min\{n; X_n(\omega) \in A\},$$

where we define $\min \emptyset = \infty$. τ_A is the time that (X_n) hits A for the first time. $\min \emptyset = \infty$ means that X never reach A and we define $\tau_A = \infty$ in this case. τ_A is a Markov time.

∵ By rewriting

$$\{\tau_A \le n\} = \{X_1 \in A\} \cup \{X_2 \in A\} \cup \cdots \cup \{X_n \in A\},$$

we see $\{\tau_A \le n\} \in \mathcal{F}_n$. □

[Theorem 9.2] For two Markov times τ and σ, set

$$\tau \vee \sigma \equiv \max\{\tau, \sigma\}, \ \tau \wedge \sigma \equiv \min\{\tau, \sigma\}, \ \tau + \sigma.$$

Then, there are all Markov times. □

∵

$$\{\tau \vee \sigma \leq n\} = \{\tau \leq n\} \cap \{\sigma \leq n\} \in \mathcal{F}_n$$

$$\{\tau \wedge \sigma \leq n\} = \{\tau \leq n\} \cup \{\sigma \leq n\} \in \mathcal{F}_n$$

$$\{\tau + \sigma \leq n\} = \bigcup_{k=1}^{n-1} \left[\{\tau = k\} \cap \{\sigma \leq n-k\}\right] \in \mathcal{F}_n \qquad \square$$

[Question asked on 3/25] Let τ, σ be $(\mathcal{F}_n)$-Markov times such that $\tau < \sigma$. Is $\sigma - \tau$ also a Markov time?

[Remark] If $\tau \leq \sigma$, $\sigma - \tau$ may take the value 0. We can actually define Markov time with values in $\{0\} \cup \mathbb{N} \cup \{\infty\}$ by adding $\mathcal{F}_0$ properly.

[Answer] $\sigma - \tau$ is $(\tilde{\mathcal{F}}_n)$-Markov time, if we define a new filtration $(\tilde{\mathcal{F}}_n)$ by $\tilde{\mathcal{F}}_n = \mathcal{F}_{n+\tau}$, which is defined as in Definition 9.3 below.

$\because$ $\{\sigma - \tau \leq n\} = \{\sigma \leq n + \tau\} \underset{(\star)}{\in} \mathcal{F}_{n+\tau} = \tilde{\mathcal{F}}_n, \ {}^\forall n.$
To show $(\star)$, by the definition of $\mathcal{F}_{n+\tau}$, we may prove

$$\{\sigma \leq n + \tau\} \cap \{n + \tau \leq k\} \in \mathcal{F}_k, \ {}^\forall k \text{ (indeed } k \geq n + 1).$$

However, $(\text{LHS}) = \cup_{j=1}^{k} \{\sigma \leq j\} \cap \{n + \tau = j\} \in \mathcal{F}_k$, since $\{\sigma \leq j\} \in \mathcal{F}_j \subset \mathcal{F}_k$ and $\{n + \tau = j\} \in \mathcal{F}_{j-n} \subset \mathcal{F}_k$. □

[Example] Let $X = (X_n)$ be $\mathbb{R}^d$-valued $(\mathcal{F}_n)$-adapted stochastic process. For $A, B \in \mathcal{B}(\mathbb{R}^d)$ such that $A \subset B$, consider the hitting times τ_A and τ_B of X to A and B. Then $\tau_B \leq \tau_A$ holds.

For a Markov time τ, we introduce a σ-filed $\mathcal{F}_\tau$ which is a collection of all information before the time τ.

[Definition 9.3] For an $(\mathcal{F}_n)$–Markov time τ, set

$$\mathcal{F}_\tau := \{A \in \mathcal{F}; A \cap \{\tau \leq n\} \in \mathcal{F}_n, {}^\forall n\}. \qquad \square$$

[Theorem 9.3] τ, σ: Markov times, $X = (X_n)$: $(\mathcal{F}_n)$-adapted. Then,

(1) $\mathcal{F}_\tau$ is a σ–field.
(2) τ is $\mathcal{F}_\tau$–measurable.
(3) $\tau \leq \sigma$ (i.e., $\tau(\omega) \leq \sigma(\omega), {}^\forall \omega \in \Omega$) $\Longrightarrow \mathcal{F}_\tau \subset \mathcal{F}_\sigma$
(4) $\tau < \infty$ (a.s.) $\Longrightarrow X_\tau$:$\mathcal{F}_\tau$-measurable

where $X_\tau(\omega) := X_{\tau(\omega)}(\omega)$ (if necessary, set $X_\infty(\omega) := 0$) $\square$

P: Assume $(\Omega, \mathcal{F}, P)$ is complete and $\mathcal{N} \subset \mathcal{F}_1$, where $\mathcal{N} = \{N \in \mathcal{F}; P(N) = 0\}$. Then, (3) holds if $\tau \leq \sigma$ a.s.

[Proof] proof of (1)

$\because$ (i) $\Omega \cap \{\tau \le n\} = \{\tau \le n\} \in \mathcal{F}_n \therefore\ \Omega \in \mathcal{F}_\tau$
(ii) For $A \in \mathcal{F}_\tau$,
$A^c \cap \{\tau \le n\} = \{\tau \le n\} \setminus (A \cap \{\tau \le n\}) \in \mathcal{F}_n \therefore\ A^c \in \mathcal{F}_\tau$
(iii) $A_k \in \mathcal{F}_\tau \Longrightarrow \bigcup_{k=1}^{\infty} A_k \in \mathcal{F}_\tau$ is easy. □

proof of (2)

$\because$ It is enough to show $\{\tau = k\} \in \mathcal{F}_\tau, {}^\forall k$. For this, we may show $\{\tau = k\} \cap \{\tau \le n\} \in \mathcal{F}_n, {}^\forall n$. However, this follows by observing

$$\{\tau = k\} \cap \{\tau \le n\} = \begin{cases} \{\tau = k\}, & \text{when } k \le n \\ \quad \emptyset \quad , & \text{when } k > n. \end{cases}$$ □

proof of (3)

$\because$ Take $^{\forall}A \in \mathcal{F}_{\tau}$. Since $\{\sigma \leq n\} \subset \{\tau \leq n\}$ by $\tau \leq \sigma$, we have

$$\begin{aligned} A \cap \{\sigma \leq n\} &= A \cap \left[\{\tau \leq n\} \cap \{\sigma \leq n\}\right] \\ &= \left[A \cap \{\tau \leq n\}\right] \cap \{\sigma \leq n\} \in \mathcal{F}_n \end{aligned}$$

$\therefore\ A \in \mathcal{F}_{\sigma}\ \therefore\ \mathcal{F}_{\tau} \subset \mathcal{F}_{\sigma}$ is shown. □

proof of (4)

$\because$ It is enough to show $\{X_{\tau} \leq a\} \in \mathcal{F}_{\tau}, ^{\forall} a \in \mathbb{R}$,
i.e., $\{X_{\tau} \leq a\} \cap \{\tau \leq n\} \in \mathcal{F}_n, ^{\forall} n$. However, this follows from $\{X_{\tau} \leq a\} \cap \{\tau \leq n\} = \cup_{k=1}^{n} \{X_k \leq a, \tau = k\}$. □

9.4 Doob's optional sampling theorem

Martingale or submartingale property is invariant for random times, if they are bounded (in ω) Markov times.

[Theorem 9.4] Let a submartingale (X_n) and Markov times τ, σ be given and assume that $\tau \leq \sigma \leq K$ (K is a constant) holds. Then, X_τ, X_σ are both integrable and

$$E[X_\sigma | \mathcal{F}_\tau] \geq X_\tau$$

holds. In particular, if (X_n) is a martingale, the above holds with equality. □

[proof] • Once this is shown for submartingale, the equality for martingale X is easy, since both $X, -X$ are submartingales.

• However, we actually show the theorem first for martingale and then apply Doob decomposition for submartingale.

• It was shown in Theorem 9.3-(4) that X_τ, X_σ are r.v.'s. The integrability of X_τ follows by

$$E[|X_\tau|] = \sum_{k=1}^{K} E[|X_k|, \tau = k] \leq \sum_{k=1}^{K} E[|X_k|] < \infty.$$

• We show the equality for martingale.

⨀ We may show $E[X_\sigma, A] = E[X_\tau, A]$, ${}^\forall A \in \mathcal{F}_\tau$. First, decompose

$$E[X_\sigma, A] = \sum_{k=1}^{K} E[X_k, A \cap \{\sigma = k\}].$$

From $\tau \le \sigma$, $A \in \mathcal{F}_\tau$ is also $A \in \mathcal{F}_\sigma$ so that $A \cap \{\sigma = k\} \in \mathcal{F}_k$. But, since (X_n) is a martingale, for ${}^\forall B \in \mathcal{F}_k$, if $1 \le k \le K$, we have

$$E[X_k, B] = E\big[E[X_K|\mathcal{F}_k], B\big] = E[X_K, B].$$

Thus, taking $B = A \cap \{\sigma = k\}$, we obtain

$$E[X_\sigma, A] = \sum_{k=1}^{K} E[X_K, A \cap \{\sigma = k\}] = E[X_K, A].$$

Similarly, $E[X_\tau, A] = E[X_K, A]$ is shown. From these, we obtain $E[X_\sigma, A] = E[X_\tau, A]$, ${}^\forall A \in \mathcal{F}_\tau$. □

- Submartingale can be decomposed as $X_n = M_n + A_n$ so that

$$E[X_\sigma|\mathcal{F}_\tau] = E[M_\sigma|\mathcal{F}_\tau] + E[A_\sigma|\mathcal{F}_\tau].$$

Since M is a martingale, we have $E[M_\sigma|\mathcal{F}_\tau] = M_\tau$ for the 1st term. For the 2nd term, since (A_n) is an increasing process and $\sigma \geq \tau$, we have $A_\sigma \geq A_\tau$.

Thus, by Theorem 9.3-(4), also noting that A_σ, A_τ are r.v.'s, we obtain

$$E[A_\sigma|\mathcal{F}_\tau] \geq E[A_\tau|\mathcal{F}_\tau] = A_\tau.$$

$\therefore$

$$E[X_\sigma|\mathcal{F}_\tau] \geq M_\tau + A_\tau = X_\tau$$

Theorem 9.4 is shown. □

[Example] Theorem 9.4 is, in general, not true for unbounded Markov times τ, σ. For example, let $(X_n)_{n=1,2,\dots}$ be a symmetric simple random walk on $\mathbb{Z}$ starting at 0
(i.e. a random motion moving one step to right or left with probability $1/2$). Set

$$\tau_{-k} = \min\{n; X_n = -k\}, \quad k = 1, 2, \dots$$

Then, (X_n) is a martingale (with a natural filtration) and $\tau_{-1} < \tau_{-2}$. Moreover, by the recurrence of random walk, $\tau_k < \infty$ *a.s.* holds. Though τ_{-1}, τ_{-2} are Markov times, since

$$X_{\tau_{-1}} = -1 > X_{\tau_{-2}} = -2 \quad a.s.,$$

Theorem 9.4 does not hold.

Random Time Change of (sub)martingale by an increasing sequence of bounded Markov times is (sub)martingale:

[Theorem 9.5] (Optional sampling theorem)
(X_n): $(\mathcal{F}_n)$-submartingale (or martingale)
$(\tau_k)_{k=1,2,\ldots}$: increasing sequence of $(\mathcal{F}_n)$–Markov times (i.e. $\tau_1 \leq \tau_2 \leq \cdots$) and each τ_k is bounded $(\sup_\omega \tau_k(\omega) < \infty, {}^\forall k)$
$\Longrightarrow (Y_k := X_{\tau_k})_{k=1,2,\ldots}$ is $(\mathcal{F}_{\tau_k})$-submartingale (or martingale) □

⊙ Theorem 9.3-(1), (3) shows that $(\mathcal{F}_{\tau_k})_{k=1,2,\ldots}$ is a filtration, while Theorem 9.3-(4) shows $(\mathcal{F}_{\tau_k})$–adaptedness of (Y_k). The integrability of Y_k and

$$E[Y_{k+1}|\mathcal{F}_{\tau_k}] \geq Y_k$$

follow from Theorem 9.4. □