Martingales and Almost Sure Convergence

김지수 (Jisu KIM)

확률론 2 (Probability Theory 2), 2025 2nd semester

Some parts of this lecture note are from the lecture notes from Prof. Alessandro Rianldo's "Advanced Probability Overview".

Martingales (마팅게일)

Martingales are elegant and powerful tools to study sequences of dependent random variables. It is originated from gambling, where a gambler can adjust the bet according to the previous results.

Definition. An increasing sequence of sub $\sigma-$fields $\{\mathcal{F}_n : \ n = 1, 2, \cdots\}$ of $\mathcal{F}$ is called a filtration. We call $\{X_n\}$ is adapted to $\{\mathcal{F}_n\}$ if $X_n \in \mathcal{F}_n$ for all n.

Definition. If $\{X_n\}$ is a sequence of random variables with

(i) $\mathbb{E}|X_n| < \infty$

(ii) $X_n \in \mathcal{F}_n$

(iii) $\mathbb{E}(X_{n+1}|\mathcal{F}_n) = X_n$ a.s. for all n

then $\{X_n\}$ is a martingale (마팅게일) (with respect to $\mathcal{F}_n$)

Definition. If (iii) is replaced with $\mathbb{E}(X_{n+1}|\mathcal{F}_n) \geq X_n$, it is a submartingale.
If (iii) is replaced with $\mathbb{E}(X_{n+1}|\mathcal{F}_n) \leq X_n$, it is a supermartingale.

Remark. If $\{X_n, \mathcal{F}_n\}$ is a submartingale, $\{-X_n, \mathcal{F}_n\}$ is a supermartingale.
If $\{\mathcal{F}_n\}$ is not specified when a martingale $\{X_n\}$ is defined, we let $\mathcal{F}_n = \sigma(X_1, \cdots, X_n)$.
If $\{\mathcal{G}_n\}$ is a filtration such that $\mathcal{G}_n \subset \mathcal{F}_n$ and $X_n \in \mathcal{G}_n$, then $\{X_n, \mathcal{G}_n\}$ is also a martingale.

We begin by describing three examples related to random walk. Let $\xi_1, \xi_2, \cdots$ be independent and identically distributed, with $\mathbb{E}\,|\xi_i| < \infty$, and let $\mu = \mathbb{E}\xi_i$. Let S_0 be a constant, and $S_n = S_0 + \xi_1 + \cdots + \xi_n$. Let $\mathcal{F}_n = \sigma(\xi_1, \ldots, \xi_n)$ for $n \geq 1$ and take $\mathcal{F}_0 = \{\emptyset, \Omega\}$.

Example ([1, Example 4.2.1])**.** Linear Martingale. If $\mu = 0$, then $\{S_n\}$ is a martingale with respect to $\mathcal{F}_n$.
To prove this, we observe that $S_n \in \mathcal{F}_n$, $\mathbb{E}|S_n| < \infty$, and ξ_{n+1} is independent of $\mathcal{F}_n$, so using the linearity of conditional expectation,

$$\mathbb{E}(S_{n+1} \mid \mathcal{F}_n) = \mathbb{E}(S_n \mid \mathcal{F}_n) + \mathbb{E}(\xi_{n+1} \mid \mathcal{F}_n) = S_n + \mathbb{E}\xi_{n+1} = S_n.$$

If $\mu \leq 0$, then $\mathbb{E}(X_{n+1} \mid \mathcal{F}_n) \leq X_n$, i.e., X_n is a supermartingale.
If $\mu \geq 0$ then S_n is a submartingale.
For any value of μ, by letting $\xi_i' = \xi_i - \mu$, we see that $S_n - n\mu$ is a martingale.

Example. Quadratic martiagale. Suppose now that $\mu = \mathbb{E}\xi_i = 0$ and $\sigma^2 = \mathrm{var}(\xi_i) < \infty$. In this case $S_n^2 - n\sigma^2$ is a martingale.

Since $(S_n + \xi_{n+1})^2 = S_n^2 + 2S_n\xi_{n+1} + \xi_{n+1}^2$ and ξ_{n+1} is independent of $\mathcal{F}_n$, we have

$$\mathbb{E}(S_{n+1}^2 - (n+1)\sigma^2 \mid \mathcal{F}_n) = S_n^2 + 2S_n\mathbb{E}(\xi_{n+1} \mid \mathcal{F}_n) + \mathbb{E}(\xi_{n+1}^2 \mid \mathcal{F}_n) - (n+1)\sigma^2 = S_n^2 - n\sigma^2.$$

Example. Exponential martingale. Let $Y_1, Y_2, \ldots$ be nonnegative i.i.d. random variables with $\mathbb{E}Y_m = 1$. If $\mathcal{F}_n = \sigma(Y_1, \ldots, Y_n)$ then

$$M_n = \prod_{m \le n} Y_m$$

defines a martingale. To prove this note that

$$\mathbb{E}(M_{n+1} \mid \mathcal{F}_n) = M_n\mathbb{E}(Y_{n+1} \mid \mathcal{F}_n) = M_n.$$

Suppose now that $Y_i = e^{\theta\xi_i}$ and $\varphi(\theta) = \mathbb{E}e^{\theta\xi_i} < \infty$. $Y_i = \exp(\theta\xi_i)/\varphi(\theta)$ has mean 1 so $\mathbb{E}Y_i = 1$ and

$$M_n = \prod_{i=1}^{n} Y_i = \frac{\exp(\theta S_n)}{\varphi(\theta)^n}$$

is a martingale.

Theorem ([1, Theorem 4.2.4, 4.2.5])**.** *(i) If X_n is a supermartingale then for $n > m$, $\mathbb{E}(X_n|\mathcal{F}_m) \le X_m$*

(ii) If X_n is a submartingale then for $n > m$, $\mathbb{E}(X_n|\mathcal{F}_m) \ge X_m$

(iii) If X_n is a martingale (마팅게일) then for $n > m$, $\mathbb{E}(X_n|\mathcal{F}_m) = X_m$

Proof. (i)

The definition gives the result for $n = m + 1$. Suppose $n = m + k$ with $k \ge 2$. By [1, Theorem 4.1.2],

$$\mathbb{E}(X_{m+k} \mid \mathcal{F}_m) = \mathbb{E}(\mathbb{E}(X_{m+k} \mid \mathcal{F}_{m+k-1}) \mid \mathcal{F}_m) \le \mathbb{E}(X_{m+k-1} \mid \mathcal{F}_m),$$

by the definition of supermartingale and [1, Theorem 4.1.9 (b)]. The desired result now follows by induction.

(ii)

Note that $-X_n$ is a supermartingale and use [1, Theorem 4.1.9 (a)].

(iii)

Observe that X_n is both a supermartingale and a submartingale. □

Remark. The idea in the proof of (ii) and (iii) will be used many times below. To keep from repeating ourselves, we will just state the result for either supermartingales or submartingales and leave it to the reader to translate the result for the other two.

Theorem ([1, Theorem 4.2.6])**.** *Let X_n be a martingale (마팅게일) with respect to $\mathcal{F}_n$ and φ be a convex function (볼록함수) with $\mathbb{E}|\varphi(X_n)| < \infty$ for all n. Then, $\varphi(X_n)$ is a submartingale with respect to $\mathcal{F}_n$. Consequently, if $p \ge 1$ and $\mathbb{E}|X_n|^p < \infty$ for all n, then $|X_n|^p$ is a submartingale with respect to $\mathcal{F}_n$.*

Proof. By Jensen's inequality and the definition,

$$\mathbb{E}(\varphi(X_{n+1}) \mid \mathcal{F}_n) \ge \varphi(\mathbb{E}(X_{n+1} \mid \mathcal{F}_n)) = \varphi(X_n).$$

□

Theorem ([1, Theorem 4.2.7])**.** *Let X_n be a submartingale with respect to $\mathcal{F}_n$, and φ be an increasing convex function(볼록증가함수) with $\mathbb{E}|\varphi(X_n)| < \infty$ for all n. Then $\varphi(X_n)$ is a submartingale with respect to $\mathcal{F}_n$. Consequently,*

(i) If X_n is a submartingale then $(X_n - a)^+$ is a submartingale.

(ii) If X_n is a supermartingale then $X_n \wedge a$ is a supermartingale.

Proof. By Jensen's inequality and the assumptions,

$$\mathbb{E}(\varphi(X_{n+1}) \mid \mathcal{F}_n) \geq \varphi(\mathbb{E}(X_{n+1} \mid \mathcal{F}_n)) \geq \varphi(X_n).$$

$\square$

Definition. A sequence of random variables $\{H_n : \ n = 1, 2, \cdots\}$ is said to be predictable relative to the filtration $\{\mathcal{F}_n\}$ if $H_n \in \mathcal{F}_{n-1}$ for all $n \geq 1$.

In words, the value of H_n may be predicted (with certainty) from the information available at time $n-1$. In this section, we will be thinking of H_n as the amount of money a gambler will bet at time n. This can be based on the outcomes at times $1, \ldots, n-1$ but not on the outcome at time n.

Once we start thinking of H_n as a gambling system, it is natural to ask how much money we would make if we used it. Let X_n be the net amount of money you would have won at time n if you had bet one dollar each time. If you bet according to a gambling system H then your winnings at time n would be as follows:

Definition. Let X_n be a (sub, super) martingale with respect to $\mathcal{F}_n$ and let $\{H_n\}$ be predictable. Define

$$(H \bullet X)_n = \sum_{m=1}^{n} H_m (X_m - X_{m-1}).$$

since if at time m you have wagered \$3 the change in your fortune would be 3 times that of a person who wagered \$1. Alternatively you can think of X_m as the value of a stock and H_m the number of shares you hold from time $m-1$ to time m.

Suppose now that $\xi_m = X_m - X_{m-1}$ have $P(\xi_m = 1) = p$ and $P(\xi_m = -1) = 1 - p$. A famous gambling system called the **martingale** is defined by $H_1 = 1$ and for $n \geq 2$,

$$H_n = \begin{cases} 2H_{n-1}, & \text{if } \xi_{n-1} = -1, \\ 1, & \text{if } \xi_{n-1} = 1. \end{cases}$$

In words, we double our bet when we lose, so that if we lose k times and then win, our net winnings will be 1. To see this consider the following concrete situation:

n	1	2	4	8	16
ξ_n	-1	-1	-1	-1	1
$(H \cdot X)_n$	-1	-3	-7	-15	1

This system seems to provide us with a "sure thing" as long as $P(\xi_m = 1) > 0$. However, the next result says there is no system for beating an unfavorable game.

Theorem. *Let X_n be a (sub, super) martingale with respect to $\mathcal{F}_n$ and let $\{H_n\}$ be predictable. Suppose $(H \bullet X)_n$ is integrable (this holds in particular when each H_n is bounded).*

Then $\{(H \bullet X)_n, \ \mathcal{F}_n\}$ is a martingale.

If $H_n \geq 0$, it is a (sub, super) martingale.

Proof. Using the fact that conditional expectation is linear, $(H \cdot X)_n \in \mathcal{F}_n$, $H_n \in \mathcal{F}_{n-1}$, and [1, Theorem 4.1.14], we have

$$\begin{aligned}\mathbb{E}((H \cdot X)_{n+1} \mid \mathcal{F}_n) &= (H \cdot X)_n + \mathbb{E}(H_{n+1}(X_{n+1} - X_n) \mid \mathcal{F}_n). \\ &= (H \cdot X)_n + H_{n+1}\mathbb{E}((X_{n+1} - X_n) \mid \mathcal{F}_n) \leq (H \cdot X)_n,\end{aligned}$$

since $\mathbb{E}((X_{n+1} - X_n) \mid \mathcal{F}_n) \leq 0$ and $H_{n+1} \geq 0$. □

Remark. $(H \bullet X)_n$ is called an integral transformation of H_n with respect to X_n, and is a discrete-time version of the famous Ito's stochastic integral
$(H \bullet X)_t = \int_0^t H_s dX_s$
or the equivalent differential equation form
$d(H \bullet X)_t = H_t dX_t$ with $(H \bullet X)_0 = 0$

We will now consider a very special gambling system: bet \$1 at each time $n \leq N$ then stop playing.

Definition. A random variable N is said to be a stopping time (정지시간) if $\{N = n\} \in \mathcal{F}_n$ for all $n < \infty$, i.e., the decision to stop at time n must be measurable with respect to the information known at that time.

Theorem ([1, Theorem 4.2.9])**.** *If N is a stopping time (정지시간) and X_n is a supermartingale, then $X_{N \wedge n}$ is a supermartingale.*

Proof. If we let $H_n = 1_{\{N \geq n\}}$, then $\{N \geq n\} = \{N \leq n-1\}^c \in \mathcal{F}_{n-1}$, so H_n is predictable. Then [1, Theorem 4.2.8] implies that

$$(H \cdot X)_n = X_{N \wedge n} - X_0$$

is a supermartingale. Since the constant sequence $Y_n = X_0$ is a supermartingale and the sum of two supermartingales is also supermartingale, we have that $X_{N \wedge n}$ is a supermartingale. □

Almost sure convergence (거의 확실한 수렴)

Submartingale is a stochastic version of monotone increasing sequence. Hence, we expect it converges a.s. if $\sup_n \mathbb{E} X_n^+ < \infty$.

Although [1, Theorem 4.2.8] implies that you cannot make money with gambling systems, you can prove theorems with them. Let $a < b$ and $\{X_n\}$ be stochastic process. Let $N_0 = 1$ and

$$\begin{aligned} N_1 &= \min\{m \geq 1 : \; X_m \leq a\} \\ N_2 &= \min\{m > N_1 : \; X_m \geq b\} \\ &\vdots \\ N_{2k-1} &= \min\{m > N_{2k-2} : X_m \leq a\} \\ N_{2k} &= \min\{m > N_{2k-1} : \; X_m \geq b\}. \end{aligned}$$

Then N_j are stopping times and

$$\{N_{2k-1} < m \leq N_{2k}\} = \{N_{2k-1} \leq m-1\} \cap \{N_{2k} \leq m-1\}^{\mathsf{C}} \in \mathcal{F}_{m-1}.$$

and so

$$H_m = \begin{cases} 1, & \text{if } N_{2k-1} < m \le N_{2k} \text{ for some } k, \\ 0, & \text{otherwise,} \end{cases}$$

defines a predictable sequence. $X_{N_{2k-1}} \le a$ and $X_{N_{2k}} \ge b$, so between times N_{2k-1} and N_{2k}, X_m crosses from below a to above b. H_m is a gambling system that tries to take advantage of these "upcrossings." In stock market terms, we buy when $X_m \le a$ and sell when $X_m \ge b$, so every time an upcrossing is completed, we make a profit of $\ge (b-a)$.

Finally, let $U_n = \sup\{k : N_{2k} \le n\}$ be the number of upcrossings completed by time n.

Theorem ([1, Theorem 4.2.10])**.** *Upcrossing inequality*

If $\{X_n\}$ is a submartingale, then

$$(b-a)\mathbb{E}(U_n) \le \mathbb{E}(X_n - a)^+ - \mathbb{E}(X_0 - a)^+.$$

(not covered in class). Let $Y_m = a + (X_m - a)^+$. By [1, Theorem 4.2.7], Y_m is a submartingale. Clearly, it upcrosses $[a, b]$ the same number of times that X_m does, and we have $(b-a)U_n \le (H \cdot Y)_n$, since each upcrossing results in a profit $\ge (b-a)$ and a final incomplete upcrossing (if there is one) makes a nonnegative contribution to the right-hand side. It is for this reason we had to replace X_m by Y_m.

Let $K_m = 1 - H_m$. Clearly, $Y_n - Y_0 = (H \cdot Y)_n + (K \cdot Y)_n$, and it follows from [1, Theorem 4.2.8] that $\mathbb{E}(K \cdot Y)_n \ge \mathbb{E}(K \cdot Y)_0 = 0$, so $\mathbb{E}(H \cdot Y)_n \le \mathbb{E}(Y_n - Y_0)$, proving the desired inequality. □

The upcrossing lemma says that a submartingale cannot cross a fixed nondegenerate interval very often with high probability. If the submartingale were to cross an interval infinitely often, then its lim sup and lim inf would have to be different and it couldn't converge.

Theorem ([1, Theorem 4.2.11])**.** *Martingale convergence theorem (마팅게일 수렴정리).*

If $\{X_n\}$ is a submartingale with $\sup_n \mathbb{E}X_n^+ < \infty$, then as $n \to \infty$, X_n converges a.s. (거의 확실한 수렴) to a limit X with $\mathbb{E}|X| < \infty$.

Proof. Since $(X-a)^+ \le X^+ + |a|$, [1, Theorem 4.2.10] implies that

$$\mathbb{E}U_n \le (|a| + \mathbb{E}X_n^+)/(b-a).$$

As $n \uparrow \infty$, $U_n \uparrow U$, where U is the number of upcrossings of $[a, b]$ by the whole sequence. So if $\sup \mathbb{E}X_n^+ < \infty$ then $\mathbb{E}U < \infty$ and hence $U < \infty$ a.s.. Since $\liminf X_n < a < b < \limsup X_n$ implies that $U = \infty$, this implies that

$$P\left(\liminf X_n < a < b < \limsup X_n\right) = 0.$$

Since this holds for all rational a and b,

$$\bigcup_{a,b \in \mathbb{Q}} \{\liminf X_n < a < b < \limsup X_n\}$$

has probability 0 and hence $\limsup X_n = \liminf X_n$ a.s., i.e., $\lim X_n$ exists a.s.. Fatou's lemma guarantees $\mathbb{E}X^+ \le \liminf \mathbb{E}X_n^+ < \infty$, so $X < \infty$ a.s. To see $X > -\infty$, we observe that

$$\mathbb{E}X_n^- = \mathbb{E}X_n^+ - \mathbb{E}X_n \le \mathbb{E}X_n^+ - \mathbb{E}X_0$$

(since X_n is a submartingale), so another application of Fatou's lemma shows

$$\mathbb{E}X^- \le \liminf_{n\to\infty} \mathbb{E}X_n^- \le \sup_n \mathbb{E}X_n^+ - \mathbb{E}X_0 < \infty.$$

□

An important special case of this theorem is:

Theorem ([1, Theorem 4.2.12])**.** *If $X_n \ge 0$ is a supermartingale, then $X_n \to X$ a.s. (거의 확실한 수렴) and $\mathbb{E}X \le \mathbb{E}X_0$.*

Proof. $Y_n = -X_n \le 0$ is a submartingale with $\mathbb{E}Y_n^+ = 0$. Since $\mathbb{E}X_0 \ge \mathbb{E}X_n$, the inequality follows from Fatou's lemma. □

We first give two "counterexamples."

Example ([1, Example 4.2.13])**.** The first shows that the assumptions of [1, Theorem 4.2.12] (or [1, Theorem 4.2.11]) do not guarantee convergence in L^1 (L^1 수렴).

Let $S_0 = 1$ and $\{S_n,\ n \ge 1\}$ be i.i.d. symmetric simple random walk. That is, $S_n = S_{n-1} + \xi_n$ where $\xi_1, \xi_2, \cdots$ are i.i.d. with $P(\xi_i = 1) = P(\xi_i = -1) = \frac{1}{2}$.

In fact, S_n does not converge and $\limsup S_n = \infty$, by [1, Exercise 5.4.1].

Let $N = \min\{n:\ S_n = 0\}$ and let $X_n = S_{N\wedge n}$. Then [1, Theorem 4.2.9] implies that X_n is a nonnegative martingale.

Hence [1, Theorem 4.2.12] implies that X_n converges to a limit X_∞ a.s. and $\mathbb{E}|X_\infty| < \infty$. In fact, $X_\infty = 0$ a.s., since convergence to $k > 0$ is impossible. (If $X_n = k > 0$ then $X_{n+1} = k \pm 1$.)

Since $\mathbb{E}X_n = \mathbb{E}X_0 = 1$ for all n and $X_\infty = 0$, $X_n \nrightarrow X$ in L_1.

The above example is an important counterexample to keep in mind as you read the rest of this chapter. The next one is not as important.

Example ([1, Example 4.2.14])**.** We will now give an example of a martingale with $X_k \to 0$ in probability (확률수렴) but not a.s. (거의 확실한 수렴).

Let $X_0 = 0$. When $X_{k-1} = 0$, let

$$X_k = \begin{cases} 1 & \text{with probability } 1/2k, \\ -1 & \text{with probability } 1/2k, \\ 0 & \text{with probability } 1 - 1/k. \end{cases}$$

When $X_{k-1} \ne 0$, let

$$X_k = \begin{cases} kX_{k-1} & \text{with probability } 1/k, \\ 0 & \text{with probability } 1 - 1/k. \end{cases}$$

From the construction, $P(X_k = 0) = 1 - 1/k$ so $X_k \to 0$ in probability.

On the other hand, the second Borel–Cantelli lemma implies $P(X_k = 0 \text{ for all } k \ge K) = 0$, and values in $(-1, 1) \setminus \{0\}$ are impossible, so X_k does not converge to 0 a.s.

References

[1] Rick Durrett. *Probability—theory and examples*, volume 49 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 2019. Fifth edition.