CHAPTER 1

Martingale Theory

We review basic facts from martingale theory. We start with discrete-time parameter martingales and proceed to explain what modifications are needed in order to extend the results from discrete-time to continuous-time. The Doob-Meyer decomposition theorem for continuous semimartingales is stated but the proof is omitted. At the end of the chapter we discuss the quadratic variation process of a local martingale, a key concept in martingale theory based stochastic analysis.

1. Conditional expectation and conditional probability

In this section, we review basic properties of conditional expectation.

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space and $\mathscr{G}$ a σ-algebra of measurable events contained in $\mathscr{F}$. Suppose that $X \in L^1(\Omega, \mathscr{F}, \mathbb{P})$, an integrable random variable. There exists a unique random variable Y which have the following two properties:

(1) $Y \in L^1(\Omega, \mathscr{G}, \mathbb{P})$, i.e., Y is measurable with respect to the σ-algebra $\mathscr{G}$ and is integrable;

(2) for any $C \in \mathscr{G}$, we have

$$\mathbb{E}\{X; C\} = \mathbb{E}\{Y; C\}.$$

This random variable Y is called the conditional expectation of X with respect to $\mathscr{G}$ and is denoted by $\mathbb{E}\{X|\mathscr{G}\}$.

The existence and uniqueness of conditional expectation is an easy consequence of the Radon-Nikodym theorem in real analysis. Define two measures on $(\Omega, \mathscr{G})$ by

$$\mu\{C\} = \mathbb{E}\{X; C\}, \quad \nu\{C\} = \mathbb{P}\{C\}, \qquad C \in \mathscr{G}.$$

It is clear that μ is absolutely continuous with respect to ν. The conditional expectation $\mathbb{E}\{X|\mathscr{G}\}$ is precisely the Radon-Nikodym derivative $d\mu/d\nu$.

If Y is another random variable, then we denote $\mathbb{E}\{X|\sigma(Y)\}$ simply by $\mathbb{E}\{X|Y\}$. Here $\sigma(Y)$ is the σ-algebra generated by Y.

The conditional probability of an event A is defined by

$$\mathbb{P}\{A|\mathscr{G}\} = \mathbb{E}\{I_A|\mathscr{G}\}.$$

The following two examples are helpful.

EXAMPLE 1.1. Suppose that the σ-algebra $\mathscr{G}$ is generated by a partition:

$$\Omega = \cup_{i=1}^{\infty} A_i, \qquad A_i \cap A_j = \emptyset \text{ if } i \neq j,$$

and $\mathbb{P}\{A_i\} > 0$. Then $\mathbb{E}\{X|\mathscr{G}\}$ is constant on each A_i and is equal to the average of X on A_i, i.e.,

$$\mathbb{E}\{X|\mathscr{G}\}(\omega) = \frac{1}{\mathbb{P}\{A_i\}} \int_{A_i} X \, d\mathbb{P}, \quad \omega \in A_i.$$

EXAMPLE 1.2. The conditional expectation $\mathbb{E}\{X|Y\}$ is measurable with respect to $\sigma(Y)$. By a well known result, it must be a Borel function $f(Y)$ of Y. We usually write symbolically

$$f(y) = \mathbb{E}\{X|Y = y\}.$$

Suppose that (X, Y) has a joint density function $p(x, y)$ on $\mathbb{R}^2$. Then we can take

$$f(y) = \frac{\int_{\mathbb{R}} x p(x, y) \, dy}{\int_{\mathbb{R}} p(x, y) \, dx}.$$

The following three properties of conditional expectation are often used.

(1) If $X \in \mathscr{G}$, then $\mathbb{E}\{X|\mathscr{G}\} = X$; more generally, if $X \in \mathscr{G}$ then

$$\mathbb{E}\{XY|\mathscr{G}\} = X\mathbb{E}\{Y|\mathscr{G}\}.$$

(2) If $\mathscr{G}_1 \subseteq \mathscr{G}_2$, then

$$\mathbb{E}\{X|\mathscr{G}_1|\mathscr{G}_2\} = \mathbb{E}\{X|\mathscr{G}_2|\mathscr{G}_1\} = \mathbb{E}\{X|\mathscr{G}_1\}.$$

(3) If X is independent of $\mathscr{G}$, then

$$\mathbb{E}\{X|\mathscr{G}\} = \mathbb{E}\{X\}.$$

The monotone convergence theorem, the dominated convergence theorem, and Fatou's lemma are three basic convergence theorems in Lebesgue integration theory. They still hold when the usual expectation is replaced by the conditional expectation with respect to an aribtary σ-algebra.

2. Martingales

A sequence $\mathscr{F}_* = \{\mathscr{F}_n, \ n \in \mathbb{Z}_+\}$ of increasing σ-algebras on Ω is called a filtration (of σ-algebras). The quadruple $(\Omega, \mathscr{F}_*, \mathscr{F}, \mathbb{P})$ with $\mathscr{F}_n \subseteq \mathscr{F}$ is a filtered probability space. We usually assume that

$$\mathscr{F} = \mathscr{F}_\infty \stackrel{def}{=} \bigvee_{n=0}^{\infty} \mathscr{F}_n,$$

the smallest σ-algebra containing all $\mathscr{F}_n$. Intuitively $\mathscr{F}_n$ represents the information of an evolving random system under consideration accumulated up to time n.

A sequence of random variables $X = \{X_n\}$ is said to be adapted to the filtration $\mathscr{F}_*$ if X_n is measurable with respect to $\mathscr{F}_n$ for all n. The filtration $\mathscr{F}_*^X$ generated by the sequence X is

$$\mathscr{F}_n^X = \sigma\{X_i, i \leq n\}.$$

It is the smallest filtration to which the sequence X is adapted.

DEFINITION 2.1. *A sequence of integrable random variables*

$$X = \{X_n,\ n \in \mathbb{Z}_+\}$$

on a filtered probability space $(\Omega, \mathscr{F}_*, \mathbb{P})$ *is called a martingale with respect to* $\mathscr{F}_*$ *if* X *is adapted to* $\mathscr{F}_*$ *and*

$$\mathbb{E}\{X_n | \mathscr{F}_{n-1}\} = X_n$$

for all n. *It is called a submartingale or supermartingale if in the above relation* $=$ *is replaced by* $\geq$ *or* $\leq$, *respectively.*

If the reference filtration is not explicitly mentioned, it is usually understood what it should be from the context. In many situations, we simply take the reference filtration is the one generated by the sequence itself.

REMARK 2.2. If X is a submartingale with respect to some filtration $\mathscr{F}_*$, then it is also a martingale with respect to its own filtration $\mathscr{F}_*^X$.

The definition of a supermartingale is rather unfortunate, for $\mathbb{E}X_n \leq \mathbb{E}X_{n-1}$, that is, the sequence of expected values is decreasing.

Intuitively a martingale represents a fair game. The defining property of a martingale can be written as

$$\mathbb{E}\{X_n - X_{n-1} | \mathscr{F}_{n-1}\} = 0.$$

We can regard the difference $X_n - X_{n-1}$ as a gambler's gain at the nth play of a game. The above equation says that even after applying all the information and knowledge he has accumulated up to time $n-1$, his expected gain is still zero.

EXAMPLE 2.3. If $\{X_n\}$ is a sequence of independent and integrable random variables with mean zero, then the partial sum

$$S_n = X_1 + X_2 + \cdots + X_n$$

is a martingale.

EXAMPLE 2.4. (1) If $\{X_n\}$ is a martingale and $f : \mathbb{R} \mapsto \mathbb{R}$ a convex function such that each $f(X_n)$ is integrable, then $\{f(X_n)\}$ is a submartingale. (2) $\{X_n\}$ is a submartingale and $f : \mathbb{R} \mapsto \mathbb{R}$ a convex and increasing function such that each $f(X_n)$ is integrable, then $\{f(X_n)\}$ is a submartingale.

EXAMPLE 2.5. (martingale transform) Let M be a martingale and $Z = \{Z_n\}$ adapted. Suppose further that each Z_n is uniformly bounded. Define

$$N_n = \sum_{i=1}^{n} Z_{i-1}(M_i - M_{i-1}).$$

Then N is a martingale. Note that in the general summand, the multiplicative factor Z_{i-1} is measurable with respect to the *left* time point of the martingale difference $M_i - M_{i-1}$.

EXAMPLE 2.6. (Reverse martingale) Suppose that $\{X_n\}$ is a sequence of i.i.d. integrable random variables and

$$Z_n = \frac{X_1 + X_2 + \cdots X_n}{n}.$$

Then $\{Z_n\}$ is a reverse martingale, which means that

$$\mathbb{E}\left\{Z_n | \mathscr{G}_{n+1}\right\} = Z_{n+1},$$

where

$$\mathscr{G}_n = \sigma\left\{S_i,\ i \geq n\right\}.$$

If we write $W_n = Z_{-n}$ and $\mathscr{H}_n = \mathscr{G}_{-n}$ for $n = -1, -2, \ldots$. Then we can write

$$\mathbb{E}\left\{W_n | \mathscr{H}_{n-1}\right\} = W_{n-1}.$$

3. Basic properties

Suppose that $X = \{X_n\}$ is a submartingale. We have $\mathbb{E}X_n \leq \mathbb{E}X_{n+1}$. Thus on average a submartingale is increasing. On the other hand, for a martingale we have $\mathbb{E}X_n = \mathbb{E}X_{n+1}$, which shows that it is purely noise. The Doob decomposition theorem claims that a submartingale can be decomposed uniquely into the sum of a martingale and an increasing sequence. The following example shows that the uniqueness question for the decomposition is not an entirely trivial matter.

EXAMPLE 3.1. Consider S_n, the sum of a sequence of independent and square integrable random variables with mean zero. Then $\left\{S_n^2\right\}$ is a submartingale. We have obviously

$$S_n^2 = S_n^2 - \mathbb{E}S_n^2 + \mathbb{E}S_n^2.$$

From $\mathbb{E}S_n^2 = \sum_{i=1}^n \mathbb{E}X_i^2$ it is easy to verify that $M_n = S_n^2 - \mathbb{E}S_n^2$ is a martingale. Therefore the above identity is a decomposition of the submartingale $\left\{S_n^2\right\}$ into the sum of a martingale and an increasing process. On the other hand,

$$S_n^2 = 2\sum_{i=1}^n S_{i-1}X_i + \sum_{i=1}^n X_i^2.$$

The first sum on the right side is a martingale (in the form of a martingale transform). Thus the above gives another such decomposition. In general these two decompositions are different.

The above example shows that in order to have a unique decomposition we need further restrictions.

DEFINITION 3.2. *A sequence $Z = \{Z_n\}$ is said to be predictable with respect to a filtration $\mathscr{F}_*$ if $\mathscr{F}_n$ if $Z_n \in \mathscr{F}_{n-1}$ for all $n \geq 0$.*

THEOREM 3.3. (Doob decomposition) *Let X be a submartingale. Then there is a unique increasing predictable process Z with $Z_0 = 0$ and a martingale M such that*

$$X_n = M_n + Z_n.$$

PROOF. Suppose that we have such a decomposition. Conditioning on $\mathscr{F}_{n-1}$ in

$$Z_n - Z_{n-1} = X_n - X_{n-1} - (M_n - M_{n-1}),$$

we have

$$Z_n - Z_{n-1} = \mathbb{E}\left\{X_n - X_{n-1}|\mathscr{F}_{n-1}\right\}.$$

The right side is nonnegative if X is a submartingale. This shows that a Doob decomposition, if exists, must be unique. It is now clear how to proceed to show the existence. We define

$$Z_n = \sum_{i=1}^{n} \mathbb{E}\left\{X_i - X_{i-1}|\mathscr{F}_{i-1}\right\}.$$

It is clear that X is increasing, predictable, and $Z_0 = 0$. Define $M_n = X_n - Z_n$. We have

$$M_n - M_{n-1} = X_n - X_{n-1} - \mathbb{E}\left\{X_n - X_{n-1}|\mathscr{F}_{n-1}\right\},$$

from which it is easy to see that $\mathbb{E}\left\{M_n - M_{n-1}|\mathscr{F}_{n-1}\right\} = 0$. This shows that M is a martingale. □

4. Optional sampling theorem

The concept of a martingale derives much of its power from the optional sampling theorem we will discuss in this section.

Most interesting events concerning a random sequence occurs not at a fixed constant time, but at a random time. The first time that the sequence reaches above a given level is a typical example.

DEFINITION 4.1. *A function $\tau : \Omega \to \mathbb{Z}_+$ on a filtered measurable space $(\omega, \mathscr{F}_*)$ is called a stopping time if $\{\tau \le n\} \in \mathscr{F}_n$ for all $n \ge 0$. Equivalently $\{\tau = n\} \in \mathscr{F}_n$ for all $n \ge 0$.*

Let X be a random sequence and τ is the first time such that $X_n > 100$. The event $\{\tau \le n\}$ means that the sequence has reached above 100 before or at time n. We can determine if $\{\tau \le n\}$ is true or false by looking at the sequence $X_1, X_2, \ldots, X_n$. We need to look at the sequence beyond time n. Therefore τ is a stopping time. On the other hand, let σ be the last time that $X_n > 100$. Knowing only the first n terms of the sequence will not determine the event $\{\sigma \le n\}$. We need to look beyond time n. Thus σ is not a stopping time.

EXAMPLE 4.2. Many years I was invited to give a talk by a fellow probabilist working in a town I had never been. When asking for the directions to the mathematics department, I was instructed to turn at the last traffic light on a street.

EXAMPLE 4.3. Let σ_n be stopping times. Then $\sigma_1 + \sigma_2$, $\sup_n \sigma_n$, and $\inf_n \sigma_n$ are all stopping times.

We have mentioned that $\mathscr{F}_n$ should be regarded as the information accumulated up to time n. We need the corresponding concept for a stopping time τ. Define

$$\mathscr{F}_\tau = \{C \in \mathscr{F}_\infty : C \cap \{\tau \le n\} \in \mathscr{F}_n \quad \text{for all} \quad n \ge 0\}.$$

It is easy to show that $\mathscr{F}_\tau$ is a σ-algebra. But more importantly, we have the following facts:

(1) $\mathscr{F}_\tau = \mathscr{F}_n$ if $\tau = n$;
(2) if $\sigma \le \tau$, then $\mathscr{F}_\sigma \subseteq \mathscr{F}_\tau$;
(3) τ is $\mathscr{F}_\tau$-measurable.
(3) if $\{X_n\}$ is adapted to $\mathscr{F}_*$, then X_τ is $\mathscr{F}_\tau$-measurable.

Note that the random variable X_τ is defined as $X_\tau(\omega) = X_{\tau(\omega)}(\omega)$.

We say that a stopping time τ is bounded if there is an integer N such that $\mathbb{P}\{\tau \le N\} = 1$. If X is a sequence of integrable random variables and τ a uniformly bounded stopping time, then X_τ is also integrable. Keep this technical point in mind when studying the following Doob's optional sampling theorem.

THEOREM 4.4. *Let X be a submartingale. Let σ and τ be two uniformly bounded stopping times such that $\sigma \le \tau$. Then*

$$\mathbb{E}\{X_\tau | \mathscr{F}_\sigma\} \ge X_\sigma.$$

PROOF. Let $C \in \mathscr{F}_\sigma$. It is enough to show that

$$\mathbb{E}\{X_\tau; C\} \ge \mathbb{E}\{X_\sigma; C\}.$$

This is implied by

$$\mathbb{E}\{X_\tau; C_n\} \ge \mathbb{E}\{X_n; C_n\}, \tag{4.1}$$

where $C_n = C \cap \{\sigma = n\}$.

For $k \ge n$, we have obviously

$$\begin{aligned}&\mathbb{E}\{X_\tau; C_n \cap (\tau = k\})\\ &\qquad = \mathbb{E}\{X_k; C_n \cap (\tau \ge k)\} - \mathbb{E}\{X_k; C_n \cap (\tau \ge k+1)\}.\end{aligned}$$

In the last term, the random variable X_k is integrated on the set $C_n(\tau \ge k+1)$. From $k \ge n$ we have $C_n \in \mathscr{F}_n \subseteq \mathscr{F}_k$, hence

$$C_n \cap (\tau \ge k+1) = C_n \cap (\tau \le k)^c \in \mathscr{F}_k.$$

Using this and the fact that X is a submartingale we have

$$\mathbb{E}\{X_k; C_n \cap (\tau \ge k+1)\} \le \mathbb{E}\{X_{k+1}; C_n \cap (\tau \ge k+1)\}.$$

It follows that

$$\begin{aligned}&\mathbb{E}\{X_\tau; C_n \cap \{\tau = k\}\}\\ &\qquad \ge \mathbb{E}\{X_k; C_n \cap (\tau \ge k)\} - \mathbb{E}\{X_{k+1}; C_n \cap (\tau \ge k+1)\}.\end{aligned}$$

We now sum over $k \geq n$. On the right side we have a telescoping sum and only the first term because τ is bounded. We also have $C_n \cap \{\tau \geq n\} = C_n$ because $\sigma \leq \tau$. It follows that $\mathbb{E}\{X_\tau; C_n\} \geq \mathbb{E}\{X_n; C_n\}$. □

COROLLARY 4.5. *Let X be a submartingale. Let $\sigma \leq \tau$ be two bounded stopping times such that $\sigma \leq \tau$. Then $\mathbb{E}\, X_\sigma \leq \mathbb{E}\, X_\tau$.*

COROLLARY 4.6. (optional sampling theorem) *Let X be a submartingale. Let $\tau_0 \leq \tau_1 \leq \tau_2 \leq \cdots$ be an increasing sequence of bounded stopping times. Then $\{X_{\tau_n}\}$ is a submartingale with respect to the filtration $\{\mathscr{F}_{\tau_n}\}$.*

COROLLARY 4.7. *Let X be a submartingale with respect to a filtration $\mathscr{F}_*$ and τ a stopping time. Then the stopped process $\{X_{n\wedge\tau}\}$ is a submartingale with respect to the original filtration $\mathscr{F}_*$.*

5. Submartingale inequalities

The L^p-norm of a random variable X is

$$\|X\|_p = \{\mathbb{E}|X|^p\}^{1/p}.$$

The set of random variables such that $\|X\|_p < \infty$ is denoted by $L^p(\Omega, \mathscr{F}, \mathbb{P})$ (or an abbreviated variant of it).

Let $X_n^* = \max_{1\leq i\leq n} X_i$. We first show that the tail probability of X_n^* is controled in some sense by the tail probability of the last element X_n.

LEMMA 5.1. (Doob's submartingale inequality) *Let X be a nonnegative submartingale. Then for any $\lambda > 0$ we have*

$$\mathbb{P}\{X_n^* \geq \lambda\} \leq \frac{1}{\lambda}\,\mathbb{E}\{X_n; X_n^* \geq \lambda\}.$$

PROOF. Let $\tau = \inf\{i : X_i \geq \lambda\}$ with the convention that $\inf \emptyset = n+1$. It is clear that τ is a stopping time; hence $\{\tau \leq n\} \in \mathscr{F}_\tau$ and by the optional sampling theorem we have

$$\mathbb{E}\{X_n; \tau \leq n\} \geq \mathbb{E}\{X_\tau; \tau \leq n\} \geq \lambda \mathbb{P}\{\tau \leq n\}.$$

In the last step we have used the fact that $X_\tau \geq \lambda$ if $\tau \leq n$. From the above inequality and the fact that $\{X_n^* \geq \lambda\} = \{\tau \leq n\}$ we have the desired inequality. □

COROLLARY 5.2. *Let X be a nonnegative submartingale. Then*

$$\mathbb{P}\{X_n^* \geq \lambda\} \leq \frac{\mathbb{E}X_n}{\lambda}.$$

EXAMPLE 5.3. The classical Kolmogorov inequality is a special case of Doob's submartingale inequality. Let $\{X_n, 1 \leq n \leq N\}$ be a sequence of independent, square integrable random variables with mean zero and $S_n = X_1 + \cdots + X_n$. Then $|S_n|^2$ is a nonnegative submartingale and we have

$$\mathbb{P}\left\{\max_{1\leq n\leq N} |S_n| \geq \lambda\right\} \leq \frac{\mathbb{E}|S_n|^2}{\lambda^2}.$$

To convert the above inequality about probability to an inequality about moments, we need the following fact about nonnegative random variables.

PROPOSITION 5.4. *Let X and Y be two nonnegative random variables. Suppose that they satisfy*

$$\mathbb{P}\left\{Y \geq \lambda\right\} \leq \frac{1}{\lambda}\,\mathbb{E}\left\{X; Y \geq \lambda\right\}$$

for all $\lambda > 0$. Then for any $p > 1$,

$$\|Y\|_p \leq \frac{p}{p-1}\,\|X\|_p.$$

For the case $p = 1$ we have

$$\|Y\|_1 \leq \frac{e}{e-1}\|1 + X\ln^+ X\|_1.$$

Here $\ln^+ x = \max\left\{\ln x, 0\right\}$ for $x \geq 0$.

PROOF. We need to truncate the random variable Y in order to deal with the possibility that the moment of Y may be infinite. Otherwise the the truncation is unnecessary and the proof may be easier to read. Let $YN = \min\left\{Y, N\right\}$. For $p > 0$, we have by Fubini's theorem,

$$\mathbb{E}Y_N^p = p\mathbb{E}\int_0^{Y_N} \lambda^{p-1}\,d\lambda = p\mathbb{E}\int_0^{\infty} \lambda^{p-1} I_{\{\lambda \leq Y_N\}} = p\int_0^N \lambda^{p-1}\mathbb{P}\left\{Y \geq \lambda\right\}\,d\lambda.$$

Now if $p > 1$, then

$$\begin{aligned}\mathbb{E}Y_N^p &= p\int_0^N \lambda^{p-1}\mathbb{P}\left\{Y \geq \lambda\right\}\,d\lambda \\ &\leq p\int_0^N \lambda^{p-2}\mathbb{E}\left\{X; Y \geq \lambda\right\}\,d\lambda \\ &= \frac{p}{p-1}\mathbb{E}\left\{XY_N^{p-1}\right\}.\end{aligned}$$

Using Hölder's inequality we have

$$\mathbb{E}Y_N^p \leq \frac{p}{p-1}\|X\|_p\,\|Y_N\|_p^{p-1}.$$

Now $\mathbb{E}Y_N^p$ is finite, hence we have above inequality

$$\|Y_N\|_p \leq \frac{p}{p1}\|X\|_p.$$

Letting $N \to \infty$ and using the monotone convergence theorem, we obtain the desired inequality follows immediately. For the proof of the case $p = 1$, see EXERCISE **??**. $\square$

The following moment inequalities for a nonnegative submartingale are often useful.

THEOREM 5.5. *Let X be a nonnegative submartingale. Then*

$$\|X_n^*\|_p \le \frac{p}{p-1}\,\|X_n\|_p, \qquad p>1;$$
$$\|X_n^*\|_1 \le \frac{e}{e-1}\|1+X_n\ln^+ X_n\|_1.$$

Here $\ln^+ x = \max\{\ln x, 0\}$.

PROOF. These inequalities follow immediately from PROPOSITION 5.4 and LEMMA 5.1. □

6. Convergence theorems

Let $X=\{X_n\}$ be a submartingale. Fix two real numbers $a<b$ and let $U_N^X[a,b]$ be the number of upcrossings of the sequence $X_1, X_2, \cdots, X_N$ from below a to above b. The precise definition of $U_N^X[a,b]$ will be clear from the proof of the next lemma.

LEMMA 6.1. (Upcrossing inequality) *We have*

$$\mathbb{E}\,U_N^X[a,b] \le \frac{\mathbb{E}(X_N-a)^+ - \mathbb{E}(X_1-a)^+}{b-a}.$$

PROOF. Let $Y_n=(X_n-a)^+$. Then Y is a positive submartingale. An upcrossing of X from below a to above b is an upcrossing of Y from 0 to $b-a$ and vice versa, so we may consider Y instead of X. Starting from $\tau_0=1$ we define the following sequence of stopping times

$$\sigma_i = \inf\{n\ge\tau_{i-1} : Y_n = 0\},$$
$$\tau_i = \inf\{n\ge\sigma_i : Y_n \ge b-a\}$$

with the convention that $\inf\emptyset = N$. If $X_{\sigma_i}\le a$ and $X_{\tau_i}\ge b$, that is, if $[\sigma_i,\tau_i]$ is an interval of a completed upcrossing from below a to above b, then we have $Y_{\tau_i}-Y_{\sigma_i}\ge b-a$. Even if $[\sigma_i,\tau_i]$ is not an interval of a completed upcrossing, we always have $Y_{\tau_i}-Y_{\sigma_i}\ge 0$. Therefore,

$$\sum_{i=1}^N (Y_{\tau_i}-Y_{\sigma_i}) \ge (b-a)U_n^X[a,b]. \tag{6.1}$$

On the other hand, since X is a submartingale and $\tau_{i-1}\le\sigma_i$, we have

$$\mathbb{E}\left[Y_{\sigma_i}-Y_{\tau_{i-1}}\right]\ge 0.$$

Now we have

$$Y_n - Y_0 = \sum_{i=1}^N \left[Y_{\tau_i}-Y_{\sigma_i}\right] + \sum_{i=1}^N\left[Y_{\sigma_i}-Y_{\tau_{i-1}}\right].$$

Taking the expected value we have

$$\mathbb{E}\left[Y_N-Y_0\right] \ge (b-a)\mathbb{E}U_N^X[a,b].$$

□

The next result is the basic convergence theorem in martingale theory.

THEOREM 6.2. *Let $\{X_n\}$ be a submartingale such that $\sup_{n\geq 1} \mathbb{E}X_n^+ \leq C$ for some constant C. Then the sequence converges with probability one to an integrable random variable X_∞.*

PROOF. We use the following easy fact from analysis: a sequence $x = \{x_n\}$ of real numbers converges to a finite number or $\pm\infty$ if and only if the number of upcrossings from a to b is finite for every pair of rational numbers $a < b$. Thus it is enough to show that

$$\mathbb{P}\left\{U_\infty^X[a,b] < \infty \text{ for all rational } a < b\right\} = 1.$$

Since the set of intervals with rational endpoints is countable, we need only to show that

$$\mathbb{P}\left\{U_\infty^X[a,b] < \infty\right\} = 1$$

for fixed $a < b$. We will show the stronger statement that $\mathbb{E}U_\infty^X[a,b] < \infty$.

From the upcrossing inequality we have

$$\mathbb{E}\, U_N^X[a,b] \leq \frac{\mathbb{E}(X_N - a)^+ - \mathbb{E}(X_1 - a)^+}{b-a}$$

for all N. From

$$(X_n - a)^+ \leq |X_n| + |a|$$

we have

$$\mathbb{E}U_N^X[a,b] \leq \frac{C + |a|}{b-a}.$$

We have $U_N^X[a,b] \uparrow U_\infty^X[a,b]$ as $N \uparrow \infty$, hence by the monotone convergence theorem, $\mathbb{E}U_N^X[a,b] \uparrow \mathbb{E}U_\infty^X[a,b]$. It follows that

$$\mathbb{E}\, U_\infty^X[a,b] \leq \frac{C + |a|}{b-a} < \infty.$$

It follows that the limit $X_\infty = \lim_{n\to} X_n$ exists with probability one.

From $|X_n| = 2X_n^+ - X_n$,

$$\mathbb{E}|X_n| \leq \mathbb{E}\left[2X_n^+ - X_n\right] \leq 2\,\mathbb{E}X_n^+ - \mathbb{E}X_1 \leq 2C - \mathbb{E}X_1.$$

By Fatou's lemma we have $\mathbb{E}|X_\infty| \leq 2C - \mathbb{E}X_1$, which shows that X_∞ is integrable. □

COROLLARY 6.3. *Every nonnegative supermartingale converges to an integrable random variable with probability one.*

PROOF. If X is a nonnegative supermartingale, then $-X$ is a nonpositive submartingale. Now apply the basic convergence THEOREM 6.2. □

7. Uniformly integrable martingales

The basic convergence THEOREM 6.2 does not claim that $\mathbb{E}X_n \to \mathbb{E}X_\infty$ or $\mathbb{E}|X_n - X_\infty| \to 0$. This does not hold in general. In many applications it is desirable to show that this convergence is in the sense of L^1. To claim the convergence in this sense, we need more conditions. One such condition that is used very often is that of uniform integrability. We recall this concept and the relevant convergence theorem.

DEFINITION 7.1. *A sequence $\{X_n\}$ is called uniformly integrable if*

$$\lim_{C\to\infty} \sup \sup \mathbb{E}\left\{|X_n|; |X_n| \geq C\right\} = 0.$$

Equivalently, for any $\epsilon > 0$, there is a constant C such that

$$\mathbb{E}\left\{|X_n|; |X_n| \geq C\right\} \leq \epsilon$$

for all n; i.e., the sequence $\{X_n\}$ has uniformly small L^1-tails.

The following equivalent condition justifies the terminology uniform integrability from another point of view.

PROPOSITION 7.2. *A sequence $\{X_n\}$ is uniformly integrable if and only if for any positive ϵ, there exists a positive δ such that*

$$\mathbb{E}\left\{|X_n|; C\right\} \leq \epsilon$$

for all n and all sets C such that $\mathbb{P}\left\{C\right\} \leq \epsilon$.

The concept of uniform integrability is useful mainly because of the following so-called Vitali convergence theorem , which supplements the three more well-known convergence theorems (Fatou, monotone, and dominated).

THEOREM 7.3. *If $X_n \to X$ in probability and $\{X_n\}$ is uniformly integrable, then X is integrable and $\mathbb{E}|X_n - X| \to 0$.*

PROOF. If is clear that the uniform integrability implies that $\mathbb{E}|X_n|$ is uniformly bounded, say $\mathbb{E}|X_n| \leq C$. There is a subsequence of $\{X_n\}$ converging to X with probability 1, hence by Fatou's lemma we have $\mathbb{E}|X| \leq C$. This shows that X is integrable. Let $Y_n = |X_n - X|$. Then $\{Y_n\}$ is also uniformly integrable and $Y_n \to 0$ in probability. We need to show that $\mathbb{E}Y_n \to 0$. For an $\epsilon > 0$ and let $C_n = \{Y_n \geq \epsilon\}$. Then $\mathbb{P}\left\{C_n\right\} \to 0$. We have

$$\mathbb{E}Y_n = \mathbb{E}\left\{Y_n; Y_n < \epsilon\right\} + \mathbb{E}\left\{Y_n; Y_n \geq \epsilon\right\}.$$

The first term is bounded by ϵ, and the second term goes to zero as $n \to \infty$ by PROPOISITION 7.2. This shows that $\mathbb{E}Y_n \to 0$. □

The following criterion for uniform integrability is often useful.

PROPOSITION 7.4. *If there is $p > 1$ and C such that $\mathbb{E}|X_n|^p < \infty$ for all n, then $\{X_n\}$ is uniformly integrable.*

PROOF. EXERCISE **??** □

EXAMPLE 7.5. Let X be an integrable random variable and $\{\mathscr{F}_n\}$ be a family of σ-algebras. Then the family of random variables $X_n = \mathbb{E}\{X_n|\mathscr{F}_n\}$ is uniformly integrable. This can be seen as follows. We may assume that X is nonnegative. Let $C_n = \{X_n \geq C\}$. Then $C_n \in \mathscr{F}_n$. We have

$$\mathbb{E}\{X_n; X_n \geq C\} = \mathbb{E}\{X; C_n\}.$$

Considering the two cases $X < K$ and $X \geq K$, we have

$$\mathbb{E}\{X_n; X_n \geq C\} \leq K\mathbb{P}\{C_n\} + \mathbb{E}\{X; X \geq K\}.$$

The second term can be made arbitrarily small for sufficiently large K. For a fixed K, the first term can be made arbitrarily small if C is sufficiently large because

$$\mathbb{P}\{C_n\} \leq \frac{\mathbb{E}X_n}{C} = \frac{\mathbb{E}X}{C}.$$

THEOREM 7.6. *If $\{X_n\}$ is a uniformly integrable submartingale, then it converges with probability 1 to an integrable random variable X_∞ and $\mathbb{E}|X_n - X_\infty| \to 0$.*

PROOF. From uniform integrability we have $\mathbb{E}|X_n| \leq C$ for some constant C; hence the submartingale converges. The L^1-convergence follows from THEOREM 7.3. □

The following theorem gives the general form of a uniformly integrable martingale.

THEOREM 7.7. *Every uniformly integrable martingale $\{X_n\}$ has the form $X_n = \mathbb{E}\{X|\mathscr{F}_n\}$ for some integrable random variable X.*

PROOF. From EXAMPLE 7.5) we see that a martingale of the indicated form is uniformly integrable. On the other hand, if the martingale $\{X_n\}$ is uniformly integrable, then $X_n \to X$ with probability 1 for an integrable random variable X. Furthermore, $\mathbb{E}|X_n - X_\infty| \to 0$. Taking the limit as $m \to \infty$ in $X_n = \mathbb{E}\{X_m|\mathscr{F}_n\}$ we have $X_n = \mathbb{E}\{X|\mathscr{F}_n\}$. □

The following theorem gives the "last term" of the uniformly integrable martingale $X_n = \mathbb{E}\{X|\mathscr{F}_n\}$.

THEOREM 7.8. *Suppose that $\mathscr{F}_*$ is a filtration of σ-algebras and that X is an integrable random variable. Then with probability 1,*

$$\mathbb{E}\{X|\mathscr{F}_n\} \to \mathbb{E}\{X|\mathscr{F}_\infty\}$$

where $\mathscr{F}_\infty = \sigma\{\mathscr{F}_n, n \geq 0\}$, the smallest σ-algebra containing all $\mathscr{F}_n, n \geq 0$.

PROOF. Let $X_n = \mathbb{E}\{X|\mathscr{F}_n\}$. Then $\{X_n, n \geq 0\}$ is a uniformly integrable martingale (see EXERCISE **??** below). Hence the limit

$$X_\infty = \lim_{n\to\infty} X_n$$

exists and the convergence is in $L^1(\Omega, \mathscr{F}, \mathbb{P})$. We need to show that $X_\infty = \{X|\mathscr{F}_\infty\}$.

It is clear that $X_\infty \in \mathscr{F}_\infty$. Letting $m \to \infty$ in $X_n = \mathbb{E}\{X_m | \mathscr{F}_n\}$, we have

$$\mathbb{E}\{X | \mathscr{F}_n\} = X_n = \mathbb{E}\{X_\infty | \mathscr{F}_n\}.$$

Now consider the collection $\mathscr{G}$ of sets $G \in \mathscr{F}_\infty$ such that

$$\mathbb{E}\{X_\infty; G\} = \mathbb{E}\{X; G\}.$$

If we can show that $\mathscr{G} = \mathscr{F}_\infty$, then by the definition of conditional expectations we have immediately $X_\infty = \mathbb{E}\{X | \mathscr{F}_\infty\}$.

The following two facts are clear:

1. $\mathscr{G}$ is a monotone class;
2. It contains every $\mathscr{F}_n$.

Therefore $\mathscr{G}$ contains the field $\cup_{n \geq 1} \mathscr{F}_n$. By the monotone class theorem (see EXERCISE **??**, $\mathscr{G}$ contains the smallest σ-algebra generated by this field. This shows that $\mathscr{F}_\infty = \mathscr{F}$, and the proof is completed. $\square$

8. Continuous time parameter martingales

Brownian motion $B = \{B_t\}$ is the most important continuous time parameter martingale. From this we will derive other examples such as $B_t^2 - t$ and $\exp[B_t - t/2]$. We will introduce Brownian motion in the next chapter. We assume that the reader has some familiarity with Brownian motion so that the topics of the next three sections are not as dry as it would have been.

The definitions of martingales, submartingales, and supermartingales extend in an obvious way to the case of continuous-time parameters. For example, a real-valued stochastic process $X = \{X_t\}$ is a submartingale with respect to a filtration $\mathscr{F}_* = \{\mathscr{F}_t\}$ if $X_t \in L^1(\Omega, \mathscr{F}_*, \mathbb{P})$ and

$$\mathbb{E}\{X_t | \mathscr{F}_s\} \geq X_s, \qquad s \leq t.$$

It is clear that for any increasing sequence $\{t_n\}$ of times the sequence $\{X_{t_n}\}$ is a submartingale with respect to $\{\mathscr{F}_{t_n}\}$. For this reason, the theory of continuous-parameter submartingales is by and large parallel to that of discrete-parameter martingales. Most of the results (except Doob's decomposition theorem) in the previous sections on discrete parameter martingales have their counterparts in continuous parameter martingales and the proofs there work in the present situation after obvious necessary modifications. However, there are two issues we need to deal with. The first one is a somewhat technical measurability problem. In order to use discrete approximations, we need to make sure that every quantity we deal with is measurable and can be approximated in some sense by the corresponding discrete counterpart. This is not true in general. For example, an expression such as

$$X_t^* = \sup_{0 \leq s \leq t} X_s \tag{8.1}$$

is in general not measurable, hence it cannot be approximated this way. The second issue is to find a continuous-time version of the Doob decomposition, which plays a central role in stochastic analysis. In Doob's theorem, we need to use the concept of predictability to ensure the uniqueness. This concept does not seem to have a straightforward generalization in the continuous-time situation. We will resolve these issues by restricting our attention to a class of submartingales satisfying some technical conditions. This class turns out to be wide enough for our applications.

The first restriction is on the filtration. We will show that this class is sufficiently rich for our applications. Suppose that $(\Omega, \mathscr{F}, \mathbb{P})$ is a probability space. A set is a null set if it is contained in a set of measure zero. Recall that a σ-algebra $\mathscr{G}$ is said to be complete (relative to $\mathscr{F}$) with respect to the probability measure $\mathbb{P}$ if it contains all null sets.

DEFINITION 8.1. *A filtration of σ-algebras $\mathscr{F}_* = \{\mathscr{F}_t,\ t \in \mathbb{R}_+\}$ on a probability space $(\Omega, \mathscr{F}_\infty, \mathbb{P})$ is said to satisfy the usual conditions if*

(1) $\mathscr{F}_0$, hence every $\mathscr{F}_t$, is complete relative to $\mathscr{F}_\infty$ with respect to $\mathbb{P}$;

(2) It is right-continuous, i.e., $\mathscr{F}_t = \mathscr{F}_{t+}$ for all $t \geq 0$, where

$$\mathscr{F}_{t+} \stackrel{\text{def}}{=} \bigcup_{s>t} \mathscr{F}_s.$$

The right side is usually denoted by $\mathscr{F}_{t+}$.

CONDITION (1) can usually be satisfied by completing every $\mathscr{F}_t$ by adding subsets of sets of measure zero. For our purpose, the most important example is the filtration generated by Brownian motion. We will show that after completion CONDITION (2) is satisfied in this case. In general these conditions have to be stated as technical assumptions. *Unless otherwise stated, from now on we will assume that all filtrations satisfy the usual condition.*

Our second restriction is on sample paths of stochastic processes.

DEFINITION 8.2. *A function $\phi : \mathbb{R}_+ \to \mathbb{R}^1$ is called a* Skorokhod *function (path) if at every point $t \in \mathbb{R}_+$ it is right continuous $\lim_{s\downarrow} \phi_s = \phi_t$ and has a finite left limit $\phi_{t-} = \lim_{s<t, s\uparrow t} \phi_s$. The space of Skorokhod functions is denote by $\mathscr{D}(\mathbb{R}_+)$ or simply $\mathscr{D}$. A stochastic process whose sample paths are Skorokhod with probability 1 is called a regular process.*

REMARK 8.3. $\mathscr{D}(\mathbb{R}_+)$ is a metric space under the so-called Skorohod metric ; see Billingsley **[2]**.

The next theorem shows that under certain not very restrictive technical conditions, every submartingale has a regular version with Skorokhod sample paths. First we clarify the meaning of a version.

DEFINITION 8.4. *We say that $\{Y_t,\ t \geq 0\}$ is a version of $\{X_t,\ t \geq 0\}$ if for all $t \geq 0$,*

$$\mathbb{P}\{X_t = Y_t\} = 1$$

It is clear from definition that if two processes are versions of each other, then they have the same family of finite-dimensional marginal distributions. For this reason, they are identical for all practical purposes. In stochastic analysis it is rarely necessary to distinguish a stochastic process from a version of it.

The following result shows that having Skorokhod paths is not a very restrictive condition for submartingales.

THEOREM 8.5. (Path regularity of submartingales) *Let that* $\{X_t,\ t \geq 0\}$ *is a submartingale with respect to a filtration* $\mathscr{F}_*$ *which satisfies the usual conditions. If the function* $t \mapsto \mathbb{E}X_t$ *is right-continuous. Then* $\{X_t,\ t \geq 0\}$ *has a regular version, i.e., there is a version of* X *almost all of whose paths are Skorokhod.*

PROOF. See Meyer [**9**] , page 95-96. □

Once the process X has Skorokhod paths, a quantity such as X_t^* defined in (8.1) becomes measurable because the maximum can be taken over a dense countable set (e.g., the set of rational or dyadic numbers). We give a proof of Doob's submartingale inequality as an example of passing from discrete to continuous-time sumartingales under the assumption of sample path regularity.

THEOREM 8.6. *(*Submartingale inequality*) Suppose that* X *is a nonnegative submartingale with regular sample paths and let* $X_T^* = \max_{0 \leq t \leq T} X_t$. *Then for any* $\lambda > 0$,

$$\lambda \mathbb{P}\left\{X_T^* \geq \lambda\right\} \leq \mathbb{E}\left\{X_T; X_T^* \geq \lambda\right\}.$$

PROOF. Without loss of generality we assume $T = 1$. Let

$$Z_N = \max_{0 \leq i \leq 2^N} X_{i/2^N}$$

for simplicity. Since the paths are regular we have $Z_N \uparrow X_1^*$. For any positive ϵ, we have for all sufficiently large N,

$$\lambda \mathbb{P}\left\{X_1^* \geq \lambda\right\} \leq \lambda \mathbb{P}\left\{Z_N \geq \lambda - \epsilon\right\} \leq \frac{\lambda}{\lambda - \epsilon} \mathbb{E}\left\{X_1; X_1 \geq \lambda - \epsilon\right\}.$$

In the last step we have used the submartingale inequality for discrete submartingales and the fact that $X_1^* \geq Z_N$. Letting $\epsilon \downarrow 0$, we obtain the desired inequality. □

THEOREM 8.7. (Submartingale convergence theorem) *Let* $X = \{X_t\}$ *be a regular submartingale. If* $\sup_{t \geq 0} \mathbb{E}X_t^+ < \infty$, *then the limit* $X_\infty = \lim_{t \to \infty} X_t$ *exists and is integrable. If* $\{X_t\}$ *is uniformly integrable, then* $\mathbb{E}|X_t - X_\infty| \to 0$.

PROOF. EXERCISE **??**. □

THEOREM 8.8. *Under the same conditions as in the preceding theorem, if in addition* $\{X_t, t \geq 0\}$ *is uniformly integrable, then the limit* X_∞ *exists a.s. Furthermore,* $\mathbb{E}|X_t - X_\infty| \to 0$ *and* $X_t = \mathbb{E}\left\{X_\infty | \mathscr{F}_t\right\}$.

PROOF. Same as in the discrete-parameter case, *mutatis mutandis*. □

9. Doob–Meyer decomposition theorem

In this section we discuss the Doob-Meyer decomposition theorem, the analogue of Doob's decomposition for continuous-time submartingales. It states, roughly speaking, that a submartingale X can be uniquely written a sum of a martingale M and an increasing process Z, i.e.,

$$X_t = M_t + Z_t.$$

This result plays a central role in stochastic analysis. There are several proofs for this theorem and they all start naturally with approximating a continuous-time submartingale by discrete-time submartingales. There are two difficulties we need to overcome. First, we need to show that these discrete-time approximations converge to the given submartingale in an appropriate sense. Second, in view of the fact in the discrete-time case, the uniqueness holds only after we assume that the increasing part of the decomposition is predictable, we need to formulate an appropriate analogue of predictability in the continuous-time case. Neither of these two problems are easy to overcome. A complete proof of this result can be found in Meyer **[9]**. The proof can be somewhat simplified under the hypothesis of continuous sample paths and the problem of predictability disappears under the usual conditions for the filtration; see Bass **[1]**. Fortunately, the Doob-Meyer theorem is one of those results for which we gain very little in going through its proof, and understanding its meaning is quite adequate for our later study. We will restrict ourselves to a form of this decomposition theorem that is sufficient for our purpose.

We always assume that the filtration $\mathscr{F}_* = \{\mathscr{F}_t,\ t \in \mathbb{R}_+\}$ satisfies the usual conditions. The definition of a martingale requires that each X_t is integrable. A local martingale is a more flexible concept.

DEFINITION 9.1. *A continuous, $\mathscr{F}_*$-adapted process $\{X_t\}$ is called a local martingale if there exists an increasng sequence of stopping times $\tau_n \uparrow \infty$ with probability one such that for each n the stopped process $X^{\tau_n} = \{X_{t\wedge\tau_n},\ t \in \mathbb{R}_+\}$ is an $\mathscr{F}_*$-martingale.*

Similarly we can define a local martingale or a local supermartingale.

EXAMPLE 9.2. Let $\{B_t\}$ be a Brownian motion in $\mathbb{R}^3$ and $a \neq 0$. We will show later that

$$X_t = \frac{1}{|B_t + a|}$$

is a local martingale.

REMARK 9.3. A word of warning is in order. For a local martingale $X = \{X_t\}$, the condition $\mathbb{E}|X_t| < \infty$ for all t does not guarantee that it is a martingale.

The following fact often useful.

PROPOSITION 9.4. *A nonnegative local martingale* $\{M_t\}$ *is a supermartingale. In particular* $\mathbb{E}M_t \leq \mathbb{E}M_0$.

PROOF. Let $\{\tau_n\}$ be the nondecreasing stopping times going to infinity such that the stopped processes M^{τ_n} are martingale. Then in particular $M_0 = M_0^{\tau_n}$ is integrable by definition. We have

$$\mathbb{E}\left\{M_{t\wedge\tau_n}|\mathscr{F}_s\right\} \leq M_{s\wedge\tau_n}.$$

We let $n \to \infty$. On the left side, we use Fatou's lemma. The limit on the right side is M_s because $\tau_n \to \infty$. Hence

$$\mathbb{E}\left\{M_t|\mathscr{F}_s\right\} \leq M_s.$$

It follows that every M_t is integrable and $\{M_t\}$ is a supermartingale. □

If with probability 1 the sample path $t \mapsto Z_t$ is an increasing function, we call Z an increasing process. If $A = Z^1 - Z^2$ for two increasing processes, then A is called a process of bounded variation.

REMARK 9.5. We will often drop the cumbersome phrase "with probability 1" when it is obviously understood from the context.

We are ready to state the important decomposition theorem.

THEOREM 9.6. (Doob-Meyer decomposition) *Suppose that the filtration* $\mathscr{F}_*$ *satisfies the usual conditions. Let* $X = \{X_t,\ t \in \mathbb{R}_+\}$ *be an* $\mathscr{F}_*$*-adapted local submartingale with continuous sample paths. There is a unique continuous* $\mathscr{F}_*$*-adapted local martingale* $\{M_t,\ t \in \mathbb{R}_+\}$ *and a continuous* $\mathscr{F}_*$*-adapted increasing process* $\{Z_t,\ t \in \mathbb{R}_+\}$ *with* $Z_0 = 0$ *such that*

$$X_t = M_t + Z_t, \quad t \in \mathbb{R}_+.$$

PROOF. See Karatzas and Shreve [**7**], page 21-30. Bass [**1**] contains a simple proof for continuous submartingales. □

In the theorem we stated that the martingale part M of a local submartingale is a local martingale. In general we cannot claim that it is a martingale even if X is a submartingale. In many applications, it is important to know that the martingale part of a submartingale is in fact a martingale so that we can claim the equality

$$\mathbb{E}X_t = \mathbb{E}M_t + \mathbb{E}Z_t.$$

We now discuss a condition for a local martingale to be a martingale, which can be useful in certain situations.

DEFINITION 9.7. *Let* $\mathscr{S}_T$ *be the set of stopping times bounded by* T*. A continuous process* $X = \{X_t\}$ *is said of* (DL)*-class if* $\{X_\sigma,\ \sigma \in \mathscr{F}_T\}$ *is uniformly integrable for all* $T \geq 0$.

This definition is useful for the following reason.

PROPOSITION 9.8. *A local martingale is a martingale if and only if it is of* (DL)*-class.*

PROOF. If M is a martingale, then

$$M_\sigma = \mathbb{E}\left\{M_T|\mathscr{F}_\sigma\right\} \quad \text{for all} \quad \sigma \in \mathscr{S}_T.$$

Hence $\{M_\sigma,\ \sigma \in \mathscr{S}_T\}$ is uniformly integrable (see EXAMPLE 7.5). Thus M is of (DL)-class.

Conversely, suppose that M is a (DL)-class local martingale and let $\{\tau_n\}$ be a sequence of stopping times going to infinity such that the stopped processes M^{τ_n} are martingales. We have

$$\mathbb{E}\left\{M_{t\wedge\tau_n}|\mathscr{F}_s\right\} = M_{s\wedge\tau_n}, \quad s \leq t.$$

As $n \to \infty$, the limit on the right side is M_s. On the left side $\{M_{t\wedge\tau_n}\}$ is uniformly integrable because $t \wedge \tau_n \in \mathscr{S}_t$. Hence we can pass to the limit on the left side and obtain $\mathbb{E}\left\{M_t|\mathscr{F}_s\right\} = M_s$. This shows that M is a martingale. □

THEOREM 9.9. *Let X be a local submartingale X and $X = M + Z$ be its Doob-Meyer decomposition. Then the following two conditions are equivalent:*

(1) *M is a martingale and Z is integrable;*
(2) *X is of (DL)-class.*

PROOF. Suppose that Z is increasing and $Z_0 = 0$. If it is integrable, i.e., $\mathbb{E}Z_t < \infty$ for all $t \geq 0$, then it is clearly of (DL)-class. On the other hand, every martingale M is of (DL)-class because the family

$$M_\sigma = \mathbb{E}\left\{M_T|\mathscr{F}_\sigma\right\},\ \sigma \in \mathscr{S}_T$$

is uniformly integrable. Thus if (1) holds, then both M and Z are of (DL)-class, hence the sum $X = M + Z$ is also of (DL)-class. This proves that (1) implies (2).

Suppose now that the submartingale X is of (DL)-class and $X = M + Z$ is its Doob-Meyer decomposition. Let $\{\tau_n\}$ be a sequence of stopping times tending to infinity such that the stopped processes M^{τ_n} are martingales. We have

$$X_{t\wedge\tau_n} = M_{t\wedge\tau_n} + Z_{t\wedge\tau_n}.$$

Taking the expectation we have

$$\mathbb{E}Z_{t\wedge\tau_n} = \mathbb{E}X_{t\wedge\tau_n} - \mathbb{E}X_0.$$

We have $Z_{t\wedge\tau_n} \uparrow Z_t$, hence $\mathbb{E}Z_{t\wedge\tau_n} \uparrow \mathbb{E}Z_t$ by the monotone convergence theorem. On the other hand, the assumption that X is of (DL)-class implies that $\mathbb{E}X_{t\wedge\tau_n} \to \mathbb{E}X_t$. Hence,

$$\mathbb{E}Z_t = \mathbb{E}X_t - \mathbb{E}X_0 < \infty.$$

Thus $\mathbb{E}Z_t$ is integrable. Since Z is also nonnegative and increasing, we see that Z is of (DL)-class. Now that both X and Z are of (DL)-class, the difference $M = X - Z$ is also of (DL)-class (see EXERCISE **??**) and is therefore a martingale. □

In these notes the broadest class of stochastic processes we will deal with is that of semimartingales.

DEFINITION 9.10. *A stochastic process is called a semimartingale (with respect to a given filtration of σ-algebras if it is the sum of a continuous local martingale and a continuous process of bounded variation.*

In real analysis it is well known that a function of bounded variation is the sum of two increasing functions. Using this result, we can show easily that a semimartingale X has the form $X = M + Z^1 - Z^2$, where M is a local martingale and Z^i are two incrasing processes.

The significance of this class of stochastic processes lies in the fact that it is closed under most common operations we usually perform on them. In particular, Itô's formula shows that a smooth function of a semimartingale is still a semimartingale.

For a semimartingale, the Doob-Meyer decomposition is still unique.

PROPOSITION 9.11. *A semimartingale X can be uniquely decomposed into the sum $X = M + Z$ of a continuous local martingale M and a continuous process of bounded variation Z with $Z_0 = 0$.*

PROOF. It is enough to show that if $M + Z = 0$ then $M = 0$ and $Z = 0$. Let $Z_t = Z_t^1 - Z_t^2$, where Z^i are increasing processes with $Z_0^1 = Z_0^2 = 0$. We have $M_t + Z_t^1 = Z_t^2$. By the uniqueness of Doob-Meyer decompositions we have $M_t = 0$ and $Z_t^1 = Z_t^2$, hence $Z_t = 0$. □

10. Quadratic variation of a continuous martingale

The quadratic variation is an important characterization of a continuous martingale. In fact Lévy's criterion shows that Brownian motion B is a martingale completely characterized by its quadratic variation $\langle B \rangle_t = t$.

Let $M = \{M_t\}$ be a continuous local martingale. Then M^2 is a continuous local submartingale. By the Doob-Meyer decomposition theorem, there is a continuous increasing process, which we will denote by $\langle M, M \rangle = \{\langle M, M \rangle_t\}$ or simply $\langle M \rangle = \{\langle M \rangle_t\}$, with $\langle M, M \rangle_0 = 0$, such that $M^2 - \langle M, M \rangle$ is a continuous local martingale. This increasing process $\langle M, M \rangle$ is uniquely determined by M and is called the quadratic variation process of the local martingale M.

EXAMPLE 10.1. The most important example is $\langle B \rangle_t = t$. This means that $B_t^2 - t$ is a (local) martingale. This can be verified directly from the definition of Brownian motion. We will show that the quadratic variation process of the martingale $B_t^2 - t$ is

$$4 \int_0^t B_s^2 \, ds.$$

Let's look at the quadratic variation process from another point of view. This also gives a good motivation for the term itself. Consider a partition of the time interval $[0t]$:

$$\Delta : \quad 0 = t_0 < t_1 < \cdots < t_n = t.$$

The mesh of the partition is the lengths of its largest interval:

$$|\Delta| = \sup_{l\geq 1} |t_l - t_{l-1}|.$$

Now let M be a continuous local martingale. We have the identity

$$M_t^2 = M_0^2 + 2\sum_{i=1}^n M_{t_{i-1}}(M_{t_i} - M_{t_{i-1}}) + \sum_{i=1}^n (M_{t_i} - M_{t_{i-1}})^2.$$

The second term on the right side looks like a Riemann sum and the third term is the quadratic variation of the process M along the partition. In CHAPTER 3 we will show by stochastic integration theory that the Riemann sum converges (in probability) to the stochastic integral

$$\int_0^t M_s\, dM_s$$

as $\Delta| \to 0$. The stochastic integral above is a local martingale. It follows that the quadratic variation along Δ also converges, and by the uniqueness of Doob-Meyer decompositions we have in probability,

$$\lim_{|\Delta|\to 0} \sum_{i=1}^n (M_{t_i} - M_{t_{i-1}})^2 = \langle M, M\rangle_t.$$

This relation justifies the name "quadratic variation process" for $\langle M, M\rangle$. We also have

$$M_t^2 = M_0^2 + 2\int_0^t M_s\, dM_s + \langle M, M\rangle_t.$$

This gives an explicit Doob-Meyer decomposition of the local submartingale M. It is also a simplest case of Itô's formula.

For a semimartingale $X = M + Z$ in its Doob-Meyer decomposition, we define $\langle Z, Z\rangle$ to be the quadratic variation of its martingale part, i.e.,

$$\langle X, X\rangle_t = \langle M, M\rangle_t.$$

Note that we still have

$$\lim_{|\Delta|\to 0} \sum_{i=1}^n (X_{t_i} - X_{t_{i-1}})^2 = \langle X, X\rangle_t. \tag{10.1}$$

To see this, letting $\Delta X_i = X_{t_i} - X_{t_{i-1}}$, we have

$$(\Delta X_i)^2 = (\Delta M_i)^2 + (2\Delta M_i + \Delta Z_i)\Delta Z_i.$$

By sample path continuity, we have

$$C(\Delta) \stackrel{\text{def}}{=} \max_{1\leq i\leq n} |2\Delta M_i + \Delta Z_i| \to 0$$

as $|\Delta| \to 0$. From this we have

$$\left|\sum_{i=1}^n (2\Delta M_i + \Delta Z_i) Z_i\right| \leq C(\Delta) \sum_{i=1}^n |\Delta Z_i| \leq C(\Delta)|Z|_t \to 0.$$

Here $|Z|_t$ denotes the total variation of $\{Z_s,\ 0 \le s \le t\}$. It follows that

$$\lim_{\Delta| \to 0} \sum_{i=1}^{n} (\Delta X_i)^2 = \lim_{\Delta| \to 0} \sum_{i=1}^{n} (\Delta M_i)^2 = \langle M, M \rangle_t.$$

This shows that (10.1) holds.

Finally we show that the product of two local martingales M and N is a semimartingale. The covariation process $\langle M, N \rangle$ is defined by polarization:

$$\langle M, N \rangle = \frac{\langle M+N, M+N \rangle - \langle M-N, M-N \rangle}{4}.$$

It is obviously a process of bounded variation.

PROPOSITION 10.2. *Let M and N be two local martingales. The process $MN - \langle M, N \rangle_t$ is a local martingale.*

PROOF. By the definition of covariation it is easy to verify that $MN - \langle M, N \rangle$ is equal to

$$\frac{(M+N)^2 - \langle M+N, M+N \rangle}{4} - \frac{(M-N)^2 - \langle M-N, M-N \rangle}{4}.$$

Both processes on the right side are local martingales, hence $MN - \langle M, N \rangle$ is also a local martingale. □

We will show later that the Doob-Meyer decomposition of MN is

$$M_t N_t = M_0 N_0 + \int_0^t M_s \, dN_s + \int_0^t N_s \, dM_s + \langle M, N \rangle_t.$$

11. First assignment

EXERCISE 1.1. Consider the Hilbert space $L^2(\Omega, \mathscr{F}, \mathbb{P})$ of square integrable random variables. The space $L^2(\Omega, \mathscr{G}, \mathbb{P})$ of $\mathscr{G}$-measurable square integrable random variables is a closed subspace. If $X \in L^2(\Omega, \mathscr{F}, \mathbb{P})$. Then $\mathbb{E}\{X|\mathscr{G}\}$ is the projection of X to $L^2(\Omega, \mathscr{G}, \mathbb{P})$.

EXERCISE 1.2. Conditioning reduces variance, i.e.,

$$\mathrm{Var}(X) \ge \mathrm{Var}\left(\mathbb{E}\{X|\mathscr{G}\}\right).$$

EXERCISE 1.3. Suppose that $X = \{X_n,\ n \ge 0\}$ is an i.i.d. integrable sequence. Let $S_n = X_1 + X_2 + \cdots X_n$. Then $\{S_n/n, n \ge 1\}$ is a reversed martingale with respect to the reversed filtration

$$\mathscr{F}^n = \sigma\{S_m,\ m \ge n\}, \quad n \in \mathbb{Z}_+.$$

EXERCISE 1.4. If a sequence X is adapted to $\mathscr{F}_*$ and τ is a stopping time, then $X_\tau \in \mathscr{F}_\tau$.

EXERCISE 1.5. Let X be a submartingale and σ and τ two bounded stopping times. Then

$$\mathbb{E}\{X_\tau|\mathscr{F}_\sigma\} \ge X_{\sigma \wedge \tau}.$$

EXERCISE 1.6. Let $\{X_n\}$ be a submartingale and $X_N^\dagger = \min_{0\le n\le N} X_n$. Then for any $\lambda > 0$, we have

$$\lambda\mathbb{P}\left[X_N^\dagger \le -\lambda\right] \le \mathbb{E}[X_N - X_1] - \mathbb{E}\left[X_N; X_N^\dagger \le -\lambda\right].$$

We also have for any $\lambda > 0$

$$\lambda\mathbb{E}\left[\max_{0\le n\le N}|X_n| \ge \lambda\right] \le \mathbb{E}|X_0| + 2\mathbb{E}|X_N|.$$

EXERCISE 1.7. If there is a C and $p > 1$ such that $\mathbb{E}|X_n|^p \le C$, then $\{X_n\}$ is uniformly integrable.

EXERCISE 1.8. Find a submartingale sequence such that $\mathbb{E}X_n^+ \le C$, but $\mathbb{E}X_n \to \mathbb{E}X_\infty$ does not hold.

EXERCISE 1.9. Let $\{X_i,\ i\in I\}$ be a uniformly integrable family of random variables. Then the family

$$\left\{X_i + X_j,\ (i,j)\in I\times I\right\}$$

is also uniformly integrable.

EXERCISE 1.10. Let M be a square integrable martingale. For any partition $\Delta = \{0 = t_0 < t_1 < \cdots < t_n = t\}$ of the interval $[0,t]$ we have

$$\mathbb{E}\sum_{j=1}^{n}\left[M_{t_j} - M_{t_{j-1}}\right]^2 = \mathbb{E}\langle M, M\rangle_t.$$