§ 4 Martingales

§ 4.1 Definition

Assuming that \( \{ \mathcal{F}_t, t \geq 0 \} \) is a collection of \( \sigma \)-fields on the same space \( \Omega \) and all \( \mathcal{F}_t \subset \mathcal{F}_s \).

The collection \( (\mathcal{F}_t, t \geq 0) \) of \( \sigma \)-fields on \( \Omega \) is called a filtration if

\[
\mathcal{F}_s \subset \mathcal{F}_t \quad \text{for all } 0 \leq s \leq t.
\]

Thus a filtration is an increasing stream of information.

If \( (\mathcal{F}_n, n = 0, 1, \ldots) \) is a sequence of \( \sigma \)-fields on \( \Omega \) and \( \mathcal{F}_n \subset \mathcal{F}_{n+1} \) for all \( n \), we call \( (\mathcal{F}_n) \) a filtration as well.

For our applications, a filtration is usually linked up with a stochastic process:

The stochastic process \( Y = (Y_t, t \geq 0) \) is said to be adapted to the filtration \( (\mathcal{F}_t, t \geq 0) \) if

\[
\sigma(Y_t) \subset \mathcal{F}_t \quad \text{for all } t \geq 0.
\]

The stochastic process \( Y \) is always adapted to the natural filtration generated by \( Y \):

\[
\mathcal{F}_t = \sigma(Y_s, s \leq t).
\]

Thus adaptedness of a stochastic process \( Y \) means that the \( Y_t \)'s do not carry more information than \( \mathcal{F}_t \).

If \( Y = (Y_n, n = 0, 1, \ldots) \) is a discrete-time process we define adaptedness in an analogous way: for a filtration \( (\mathcal{F}_n, n = 0, 1, \ldots) \) we require that \( \sigma(Y_n) \subset \mathcal{F}_n \).
Example. (Examples of adapted processes)
Let \((B_t, t \geq 0)\) be Brownian motion and \((\mathcal{F}_t, t \geq 0)\) be the corresponding natural filtration. Stochastic processes of the form

\[ X_t = f(t, B_t), \quad t \geq 0, \]

where \(f\) is a function of two variables, are adapted to \((\mathcal{F}_t, t \geq 0)\). This includes the processes

\[ X_t^{(1)} = B_t, \quad X_t^{(2)} = B_t^2, \quad X_t^{(3)} = B_t^2 - t, \quad X_t^{(4)} = B_t^3, \quad X_t^{(5)} = B_t^4. \]

But also processes, which may depend on the whole past of Brownian motion, can be adapted. For example,

\[ X_t^{(6)} = \max_{0 \leq s \leq t} B_s \quad \text{or} \quad X_t^{(7)} = \min_{0 \leq s \leq t} B_s^2. \]

If the stochastic process \(Y\) is adapted to the natural Brownian filtration \((\mathcal{F}_t, t \geq 0)\), we will say that \(Y\) is adapted to Brownian motion. This means that \(Y_t\) is a function of \(B_s, s \leq t\).

The following processes are not adapted to Brownian motion:

\[ X_t^{(8)} = B_{t+1}, \quad X_t^{(9)} = B_T - B_t, \quad X_t^{(10)} = B_t + B_T, \]

where \(T > 0\) is a fixed number.

- Some filtrations are bigger than others.
  Suppose \((Y_t, t \geq 0)\) represents the stock price of a particular company. \((X_t, t \geq 0)\) represents the exchange rate. Then

\[ \mathcal{F}_t = \sigma(X_s, Y_s, s \leq t) \]

is a bigger filtration than \(\sigma(Y_s, s \leq t)\).
Suppose \( \mathcal{F}_t \) represent our information up to time \( t \). We already learn that for a future time \( t' \),

\[
E[X_{t'} | \mathcal{F}_t]
\]

is our best prediction of \( X_{t'} \) given the information \( \mathcal{F}_t \). If \( X_t = B_t \), \( \sigma_{\mathcal{F}_t} = \sigma(X_s, s \leq t) \), then

\[
E[X_{t'} | \mathcal{F}_t] = X_t \quad \text{and} \quad E[B_{t'}^2 - t' | \mathcal{F}_t] = B_t^2 - t.
\]

If \( \sigma_{\mathcal{F}_t} \) is not \( \sigma(X_s, s \leq t) \), the above equation may not hold.

---

The stochastic process \( X = (X_t, t \geq 0) \) is called a continuous-time martingale with respect to the filtration \( (\mathcal{F}_t, t \geq 0) \), we write \( (X, (\mathcal{F}_t)) \), if

- \( E|X_t| < \infty \) for all \( t \geq 0 \).
- \( X \) is adapted to \( (\mathcal{F}_t) \); see p. 77 for the definition.
- 

\[
E(X_t | \mathcal{F}_s) = X_s \quad \text{for all } 0 \leq s < t,
\]

i.e. \( X_s \) is the best prediction of \( X_t \) given \( \mathcal{F}_s \).
The stochastic process $X = (X_n, n = 0, 1, \ldots)$ is called a discrete-time martingale with respect to the filtration $(\mathcal{F}_n, n = 0, 1, \ldots)$, we write $(X, (\mathcal{F}_n))$, if

- $E|X_n| < \infty$ for all $n = 0, 1, \ldots$,
- $X$ is adapted to $(\mathcal{F}_n)$,
- $E(X_{n+1} | \mathcal{F}_n) = X_n$ for all $n = 0, 1, \ldots$ \hspace{1cm} (1.43)

i.e. $X_n$ is the best prediction of $X_{n+1}$ given $\mathcal{F}_n$.

It is not difficult to see that the defining property (1.43) can be rewritten in the form

$$E(Y_{n+1} | \mathcal{F}_n) = 0, \text{ where } Y_{n+1} = X_{n+1} - X_n, \quad n = 0, 1, \ldots \hspace{1cm} (1.44)$$

It is obvious that

$$EX_s = EX_t \text{ for all } s \text{ and } t$$

is $X = (X_t, t \geq 0)$ is a martingale.

§ 4.2 Examples

**Example** Let $(Z_n)$ be a sequence of independent random variables with finite expectation and $Z_0$.

$E[Z_n] = 0$ for $n \geq 1$. Let

$$R_n = Z_0 + \ldots + Z_n, \quad n \geq 0$$

$$\mathcal{F}_n = \sigma(Z_0, Z_1, \ldots, Z_n).$$

Then $(R_n, n=0,1,2,\ldots)$ is a martingale with respect to $(\mathcal{F}_n, n=0,1,\ldots)$.
Example (Collecting information about a random variable)
Let \( Z \) be a random variable on \( \Omega \) with \( E|Z| < \infty \) and \( (\mathcal{F}_t, t \geq 0) \) be a filtration on \( \Omega \). Define the stochastic process \( X \) as follows:
\[
X_t = E(Z | \mathcal{F}_t), \quad t \geq 0.
\]

Since \( \mathcal{F}_t \) increases when time goes by, \( X_t \) gives us more and more information about the random variable \( Z \). In particular, if \( \sigma(Z) \subset \mathcal{F}_t \) for some \( t \), then \( X_t = Z \). We show that \( X \) is a martingale.

An appeal to Jensen’s inequality (A.2) on p. 188 and to Rule 2 on p. 70 yields
\[
E|X_t| = E[E(Z | \mathcal{F}_t)] \leq E[E(|Z| | \mathcal{F}_t)] = E|Z| < \infty.
\]

Moreover, \( X_t \) is obtained by conditioning on the information \( \mathcal{F}_t \). Hence it does not contain more information than \( \mathcal{F}_t \), so \( \sigma(X_t) \subset \mathcal{F}_t \). It remains to check (1.41). Let \( s < t \). Then an application of Rule 6 on p. 72 yields
\[
E(X_t | \mathcal{F}_s) = E[E(Z | \mathcal{F}_t) | \mathcal{F}_s] = E(Z | \mathcal{F}_s) = X_s.
\]

Thus \( X \) obeys the defining properties of a continuous-time martingale; see p. 80. \( \square \)

Example (Brownian motion is a martingale)
Let \( B = (B_t, t \geq 0) \) be Brownian motion. We conclude from Examples 1.4.14 and 1.4.15 that both, \( (B_t, t \geq 0) \) and \( (B^2_t - t, t \geq 0) \), are martingales with respect to the natural filtration \( \mathcal{F}_t = \sigma(B_s, s \leq t) \).

In the same way, you can show (do it!) that \( ((B^3_t - 3tB_t), (\mathcal{F}_t)) \) is a martingale.

Try to find a stochastic process \( (A_t) \) such that \( ((B^4_t + A_t), (\mathcal{F}_t)) \) is a martingale.

Hint: first calculate \( E[((B_t - B_s) + B_s)^4 | \mathcal{F}_s] \) for \( s < t \). \( \square \)

The following example of a martingale transform is a first step toward the definition of the Itô stochastic integral. Indeed, such a transform can be considered as a discrete analogue of a stochastic integral.

Example (Martingale transform)
Let \( Y = (Y_n, n = 0, 1, \ldots) \) be a martingale difference sequence with respect to the filtration \( (\mathcal{F}_n, n = 0, 1, \ldots) \); see (1.44) for the definition. Consider a stochastic process \( C = (C_n, n = 1, 2, \ldots) \) and assume that, for every \( n \), the information carried by \( C_n \) is contained in \( \mathcal{F}_{n-1} \), i.e.
\[
\sigma(C_n) \subset \mathcal{F}_{n-1}.
\] (1.45)

This means that, given \( \mathcal{F}_{n-1} \), we completely know \( C_n \) at time \( n - 1 \).
If the sequence \((\bar{C}_n, n = 1, 2, \ldots)\) satisfies (1.45), we call it \textit{previsible} or \textit{predictable} with respect to \((\mathcal{F}_n)\).

Now define the stochastic process

\[
X_0 = 0, \quad X_n = \sum_{i=1}^{n} C_i Y_i, \quad n \geq 1. \tag{1.46}
\]

For obvious reasons, the process \(X\) is denoted by \(C \bullet Y\). It is called the \textit{martingale transform of } \(Y\) \textit{by } \(C\).

The martingale transform \(C \bullet Y\) is a martingale if \(EC_n^2 < \infty\) and \(EY_n^2 < \infty\) for all \(n\). We check the three defining properties on p. 80:

\[
E|X_n| \leq \sum_{i=1}^{n} E|C_i Y_i| \leq \sum_{i=1}^{n} [E C_i^2 EY_i^2]^{1/2} < \infty,
\]

where we made use of the \textit{Cauchy–Schwarz inequality} \(E|C_i Y_i| \leq [E C_i^2 EY_i^2]^{1/2}\) (see p. 188). Clearly, \(X_n\) is adapted to \(\mathcal{F}_n\) since \(Y_1, \ldots, Y_n\) do not carry more information than \(\mathcal{F}_n\), and \(C_1, \ldots, C_n\) is predictable. Hence \(\sigma(X_n) \subset \mathcal{F}_n\). Moreover, applying Rule 5 on p. 71 and recalling that \((C_n)\) is predictable, we obtain

\[
E(X_n - X_{n-1} | \mathcal{F}_{n-1}) = E(C_n Y_n | \mathcal{F}_{n-1}) = C_n E(Y_n | \mathcal{F}_{n-1}) = 0.
\]

In the last step we used the defining property (1.44) of the martingale difference sequence \((Y_n)\). Hence \((X_n - X_{n-1})\) is a martingale difference sequence, and \((X_n)\) is a martingale with respect to \((\mathcal{F}_n)\).

\textbf{Example} \hspace{1cm} (A Brownian martingale transform)

Consider Brownian motion \(B = (B_s, s \leq t)\) and a partition

\[
0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = t.
\]

Using the independent increment property of \(B\), it is not difficult to see (check it!) that the sequence

\[
\Delta B: \quad \Delta_0 B = 0, \quad \Delta_i B = B_{t_i} - B_{t_{i-1}}, \quad i = 1, \ldots, n,
\]

forms a martingale difference sequence \(\Delta B\) with respect to the filtration given by

\[
\mathcal{F}_0 = \{\emptyset, \Omega\}, \quad \mathcal{F}_i = \sigma(B_{t_j}, 1 \leq j \leq i), \quad i = 1, \ldots, n.
\]
Now consider the transforming sequence
\[ \tilde{B} = (B_{t_{i-1}}, i = 1, \ldots, n). \]

It is predictable with respect to \((\mathcal{F}_n)\) (why?). The martingale transform \(\tilde{B} \cdot \Delta B\) is then a martingale:
\[
(\tilde{B} \cdot \Delta B)_k = \sum_{i=1}^{k} \tilde{B}_i \Delta_i B = \sum_{i=1}^{k} B_{t_{i-1}} (B_{t_i} - B_{t_{i-1}}), \quad k = 1, \ldots, n.
\]

The sums on the right-hand side have the typical form of Riemann–Stieltjes sums which would be used for the definition of the Riemann–Stieltjes integral \(\int_0^t B_s dB_s\); see Section 2.1.2. However, this integral does not exist in the Riemann–Stieltjes sense since the sample paths of Brownian motion are too irregular. We will see in Section 2.2.1 that \(\tilde{B} \cdot \Delta B\) is a discrete-time analogue of the Itô stochastic integral \(\int_0^t B_s dB_s\). 

\[\square\]

§ 4.3 The Interpretation of a Martingale as a Fair Game.

If \(X_t\) is the value of a game at time \(t\), then at time \(s\), the best prediction of the net winnings given the information \(\mathcal{F}_s\) \(s \leq t\) has value
\[
E(X_t - X_s | \mathcal{F}_s) = E(X_t | \mathcal{F}_s) - X_s
\]

If \((X_t, \mathcal{F}_t)\) is a martingale, the right-hand side is zero. This is exactly what you would expect to be a fair game.
We know that the martingale transform \( C \cdot Y \) of a martingale difference sequence \( Y \) is a martingale; see Example 1.5.6. It has an interesting interpretation in the context of fair games: think of \( Y_n \) as your net winnings per unit stake at the \( n \)th game which are adapted to a filtration \( (\mathcal{F}_n) \). Your stakes \( C_n \) constitute a predictable sequence with respect to \( (\mathcal{F}_n) \), i.e. at the \( n \)th game your stake \( C_n \) does not contain more information than \( \mathcal{F}_{n-1} \) does. At time \( n - 1 \) this is the best information we have about the game. The martingale transform \( C \cdot Y \) gives you the net winnings per game. In particular, \( (C \cdot Y)_n = \sum_{i=1}^{n} C_i Y_i \) are the net winnings up to time \( n \), and \( C_n Y_n \) are the net winnings per stake \( C_n \) at the \( n \)th game. It is fair since the best prediction of the net winnings \( C_n Y_n \) of the \( n \)th game, just before the \( n \)th game starts, is zero: \( E(C_n Y_n | \mathcal{F}_{n-1}) = 0 \).

In Chapter 2 we will learn about the continuous-time analogue to martingale transforms: the Itô stochastic integral.

**Notes and Comments**

Martingales constitute an important class of stochastic processes. The theory of martingales is considered in all modern textbooks on stochastic processes. See Karatzas and Shreve (1988) and Revuz and Yor (1991) for the advanced theory of continuous-time martingales. The book by Williams (1991) contains an introduction to conditional expectations and discrete-time martingales.