Chapter 4

Markov and Strong Markov Property

4.1 Basic Definitions of a Markov Process

Intuitively, a process $\{X_t, \mathcal{F}_t\}$ is said to be Markov if given the present state, X_s, any information about the history $\mathcal{F}_s$ of the process before time s is irrelevant for determining the evolution of X after time s. The Markov property is a property of the distribution induced by the process on the path space, and is not a sample path property. Therefore, when talking about Markov processes, we can (and will) assume without loss of generality that we are dealing with the canonical version of the Markov process, i.e., Ω is the appropriate space of functions (continuous functions in the case of Brownian motion) and $X_t(\omega) = \omega_t$ and $\mathcal{F}_t$ is the σ-algebra generated by the coordinate processes for $s \leq t$. A rigorous definition of a Markov process is given as follows.

Definition 4.1.1. *Given $d \in \mathbb{N}$ and μ a probabiity measure on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, an adapted, d-dimensional process $\{X_t, \mathcal{F}_t\}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P}^\mu)$ is said to be a* Markov process with initial distribution μ *if*

1. $\mathbb{P}^\mu(X_0 \in A) = \mu(A)$ *for every* $A \in \mathcal{B}(\mathbb{R}^d)$;
2. *for* $s, t \geq 0$ *and* $A \in \mathcal{B}(\mathbb{R}^d)$,

$$\mathbb{P}^\mu(X_{t+s} \in A | \mathcal{F}_s) = \mathbb{P}^\mu(X_{t+s} \in A | X_s), \quad \mathbb{P}^\mu \quad a.s.$$

Remark 4.1.2. Note that when $\mathcal{F}_s = \sigma(X_u, u \leq s)$, it is easy to show that the second condition above is equivalent to the condition that for every $0 \leq s_1 \leq s_2 \leq \ldots \leq s_n \leq s$, A, A_i open sets in $\mathbb{R}^d$, $i = 1, \ldots, n$ and $n \in \mathbb{N}$,

$$\mathbb{P}^\mu(X_{t+s} \in A, X_{s_1} \in A_1, X_{s_2} \in A_2, \ldots, X_{s_n} \in A_n) = \mathbb{P}^\mu(X_{t+s} \in A, X_{s_n} \in A_n).$$

When dealing with Markov processes, one is sometimes interested in properties of the entire paths after a certain time (for example the future supremum), conditioned on the past. For this, it is convenient to introduce a family of shift operators. For $s \geq 0$, let $\theta_s : \Omega \mapsto \Omega$ be a $\mathcal{F} \setminus \mathcal{F}$-measurable mapping such that

$$X_{s+t}(\omega) = X_t(\theta_s \omega).$$

For example, when Ω is the space $\mathcal{C}[0,\infty)$ of continuous functions on $[0,\infty)$ and X_t is the co-ordinate mapping process $X_t(\omega) = \omega(t)$, we can define

$$(\theta_s \omega)(t) = \omega(s+t), \quad t \geq 0.$$

If $Y : \mathcal{C}[0,\infty) \mapsto \mathbb{R}$ is $\mathcal{B}(\mathcal{C}[0,\infty))$ measurable, then $Y \circ \theta_s$ is a function of the future of the path, after time s. Indeed, consider the example $Y(\omega) = f(\omega(t))$ for some $t \in [0,\infty)$ and $f : \mathbb{R} \mapsto \mathbb{R}$. Then

$$Y \circ \theta_s = f(\theta_s(\omega(t))) = f(\omega(s+t)) = f(X_{s+t}).$$

Likewise, if $Y(\omega) = f(w_{t_1}, \ldots, w_{t_n})$ for some fixed sequence of times $t_1 < t_2 < \ldots t_n$ and $f : \mathbb{R}^n \mapsto \mathbb{R}$, then

$$Y \circ \theta_s = f(X_{s+t_1}, \ldots, (X_{s+t_n}).$$

In terms of the shift operator, the second condition of the Markov process can be shown to be equivalent to the condition

$$\mathbb{P}^x(\theta_s^{-1} A | \mathcal{F}_s) = \mathbb{P}^{X_s}(A) \qquad \mathbb{P}^x - a.s. \tag{4.1}$$

for every $A \in \mathcal{F}_\infty^X$, or equivalent to the condition that for all $s \geq 0$, and all bounded, measurable $Y : \Omega \mapsto \mathbb{R}$,

$$\mathbb{P}^x(Y \circ \theta_s | \mathcal{F}_s) = \mathbb{E}_{X_s}[Y], \tag{4.2}$$

where the right-hand side is equal to $\psi(X_s)$, where $\psi(y) \doteq \mathbb{E}_y[Y]$, $y \in \mathbb{R}$. The forms (4.1) and (4.2) are very important in practice.

4.2 Brownian Motion and the Markov Property

In this section, we show that Brownian motion $\{B_t, \mathcal{F}_t\}$ is a Markov process. The basic intuition behind why this is the case is seen by representing, for $0 \leq s \leq t$, $B_t = B_t - B_s + B_s$ and noting that $B_t - B_s$ is independent of $\mathcal{F}_s$. Thus it is clear that B_t depends on $\mathcal{F}_s$ only through B_s. To make this rigorous, we establish a basic lemma. Subsequently, making the identification $X \mapsto B_t - B_s$, $Y \mapsto B_t$, $\mathcal{F} \mapsto \mathcal{F}_s$ and $\mathcal{G} \mapsto \sigma(B_s)$, we will be able to deduce the Markov property of Brownian motion as a corollary. In fact, it is straightforward to deduce from this that, for every $s \geq 0$, $B_{s+\cdot} - B_s$ is a Brownian motion independent of $\mathcal{F}_s$.

Lemma 4.2.1. *Suppose that X and Y are d-dimensional random vectors on $(\Omega, \mathcal{F}, \mathbb{P})$, $\mathcal{G}$ is a sub-σ-algebra of $\mathcal{F}$, X is independent of $\mathcal{G}$ and Y is $\mathcal{G}$-measurable. Then for every $\Gamma \in \mathcal{B}(\mathbb{R}^d)$,*

$$\mathbb{P}(X + Y \in \Gamma \mid \mathcal{G}) = \mathbb{P}(X + Y \in \Gamma \mid Y) \quad \mathbb{P}\text{-}a.s.$$

and

$$\mathbb{P}(X + Y \in \Gamma \mid Y = y) = \mathbb{P}(X + Y \in \Gamma) \quad a.e.\ for\ \mathbb{P}\,Y^{-1}.$$

Proof. We will show that for $D \in \mathcal{B}(\mathbb{R}^d \times \mathbb{R}^d)$,

$$\mathbb{P}((X, Y) \in D \mid \mathcal{G}) = \mathbb{P}((X, Y) \in D \mid Y).$$

First look at $D = D_1 \times D_2$ for $D_1, D_2 \in \mathcal{B}(\mathbb{R}^d)$ (so-called measurable rectangles in $(\mathbb{R}^d)^2$. The left hand side is

$$\mathbb{P}(X \in D_1, Y \in D_2 \mid \mathcal{G}) = \mathbf{1}_{\{Y \in D_2\}}\, \mathbb{P}(X \in D_1 \mid \mathcal{G}) = \mathbf{1}_{\{Y \in D_2\}}\, \mathbb{P}(X \in D_1)$$

and the right hand side is equal to the same thing by the same logic. Since the measurable rectangles form a Dynkin system and generate $\mathcal{B}(\mathbb{R}^d \times \mathbb{R}^d)$, we are done. $\square$

We will now, in fact, establish a stronger version of the above result, that shows that the above result holds when deterministic times are replaced by (a.s. finite) stopping times.

Theorem 4.2.2. *Let $\{B_t, \mathcal{F}_t\}$ be a Brownian motion, where $\{\mathcal{F}_t\}$ is a right-continuous filtration, and let T be a finite-valued stopping time. Then the process $B^{(T)} = B_{T+t} - B_T$ for $t \geq 0$ is a Brownian motion independent of $\mathcal{F}_T$.*

Before we prove Theorem 4.2.2, we will establish a simple approximation lemma.

Definition 4.2.3. *A* simple stopping time *is a stopping time whose image is countable.*

Lemma 4.2.4. *Given a right-continuous filtration $\{\mathcal{F}_t\}$ and any finite $\{\mathcal{F}_t\}$-stopping time, there exists a non-increasing sequence of simple $\{\mathcal{F}_t\}$-stopping times $T_1 \geq T_2 \geq \cdots$ such that $\lim_{n\to\infty} T_n(\omega) = T(\omega)$ for every $\omega \in \Omega$ and, in addition, $\mathcal{F}_T = \bigcap_n \mathcal{F}_{T_n}$.*

Proof. For $n \in \mathbb{N}$, define for each $\omega \in \Omega$,

$$T_n(\omega) \doteq \sum_{k=0}^{\infty} \frac{k+1}{2^n} \mathbf{1}_{[k2^{-n},(k+1)2^{-n})}(T(\omega)).$$

Then clearly $T_n(\omega) \geq T_{n+1}(\omega)$ for all n, and $T_n(\omega)$ converges to $T(\omega)$ as $n \to \infty$ since $0 \leq T_n(\omega) - T(\omega) \leq 2^{-n}$. Also, for any $t \geq 0$ and $n \in \mathbb{N}$, there exists $k \in \mathbb{N}$ such that $t \in [k2^{-n}, (k+1)2^{-n})$, and

$$\{T_n \leq t\} = \{T_n \leq k2^{-n}\} = \{T \leq k2^{-n}\} \in \mathcal{F}_{k2^{-n}} \subseteq \mathcal{F}_t.$$

Therefore each T_n is a stopping time.

Now, since $T \le T_n$, it follows that $\mathcal{F}_T \subseteq \mathcal{F}_{T_n}$, and therefore $\mathcal{F}_T \subseteq \bigcap_n \mathcal{F}_{T_n}$. On the other hand, if $A \in \bigcap_n \mathcal{F}_{T_n}$ then $A \cap \{T_n \le t\} \in \mathcal{F}_t$ for all $n \ge 1$ and all $t \ge 0$. Therefore

$$A \cap \{T \le t\} = \bigcap_{\varepsilon>0} \bigcup_{m\ge1} \bigcap_{n\ge m} (A \cap \{T_n \le t+\varepsilon\}) \in \bigcap_{\varepsilon>0} \mathcal{F}_{t+\varepsilon} = \mathcal{F}_t$$

by the right continuity of $\{\mathcal{F}_t\}$. This completes the proof that $\bigcap_n \mathcal{F}_{T_n} = \mathcal{F}_T$. $\square$

Proof of Theorem 4.2.2. Step 1. We first prove the theorem for the case when T is a simple stopping time with range $\{\tau_1, \tau_2, \dots\} \subset [0,\infty)$. For any $A \in \mathcal{F}_T$ and for all $C_1, \dots, C_m \in \mathcal{B}(\mathbb{R})$, $0 \le t_1 < t_2 < \dots, < t_m < \infty$, $m \in \mathbb{N}$,

$$\begin{aligned}
&\mathbb{P}\left(A \cap \left(\textstyle\bigcap_{i\le m}\{B_{T+t_i} - B_T \in C_i\}\right)\right) \\
&\quad= \sum_{k=0}^{\infty} \mathbb{P}\left(A \cap \left(\bigcap_{i\le m}\{B_{\tau_k+t_i} - B_{\tau_k} \in C_i, T = \tau_k\}\right)\right) \\
&\quad= \sum_{k=0}^{\infty} \mathbb{P}\left(\bigcap_{i\le m}\{B_{\tau_k+t_i} - B_{\tau_k} \in C_i\}\right) \mathbb{P}\left(\{T=\tau_k\} \cap A\right) \\
&\quad= \mathbb{P}\left(\bigcap_{i\le m}\{B_{t_i} \in C_i\}\right) \mathbb{P}(A),
\end{aligned}$$

where the second equality uses the fact that $A \cap \{T = \tau_k\} \in \mathcal{F}_{\tau_k}$ along with the independence of $B_{\tau_k+\cdot} - B_{\tau_k}$ from $\mathcal{F}_{\tau_k}$, and the last equality follows from the time-homogeneity of Brownian motion. Now set $A = \mathbb{R}$ to deduce that $t \mapsto B_{T+t} - B_T$ has the same finite-dimensional distributions as Brownian motion. Since the paths $t \mapsto B_{T+t} - B_T$ are also easily seen to be continuous, it follows that $B_{T+\cdot} - B_T$ is a Brownian motion. Finally, since $A \in \mathcal{F}_T$ was arbitrary, we also conclude that $B_{T+\cdot} - B_T$ is independent of F_T.

Step 2. For a general stopping time T, consider the approximating sequence of simple stopping times $T_n \searrow T$ defined above. For any $A \in \mathcal{F}_T$ and for all open $C_1, \dots, C_m \in \mathcal{B}(\mathbb{R})$, $0 \le t_1 < t_2 < \dots, < t_m < \infty$, $m \in \mathbb{N}$, we have by the continuity of Brownian motion,

$$\begin{aligned}
\mathbb{P}\left(A \cap \left(\bigcap_{i\le m}\{B_{T+t_i} - B_T \in C_i\}\right)\right) &= \lim_{n\to\infty} \mathbb{P}\left(A \cap \left(\bigcap_{i\le m}\{B_{T_n+t_i} - B_{T_n} \in C_i\}\right)\right) \\
&= \lim_{n\to\infty} \mathbb{P}\left(\bigcap_{i\le m}\{B_{t_i} \in C_i\}\right) \mathbb{P}(A) \\
&= \mathbb{P}\left(\bigcap_{i\le m}\{B_{t_i} \in C_i\}\right) \mathbb{P}(A)
\end{aligned}$$

since $A \in \mathcal{F}_{T_n}$ for all n. $\square$