Markov Processes

Date: 11/16/2021  Scribe: Partha Dey (psdey)

In last class we sketched the proof of the following result.

**Theorem 25.1.** Let \( b, \sigma \) be globally Lipschitz in the space variable. There exists a mapping \( F : \mathbb{R}^d \times C(\mathbb{R}^+; \mathbb{R}^m) \to C(\mathbb{R}^+; \mathbb{R}^d) \) such that

1. for every \( x \in \mathbb{R}^d \), \( F(x, \cdot) \) is a measurable map. Moreover, for every \( x \in \mathbb{R}^d, t \geq 0 \), \( F(x, w)_t \) coincides \( W(dw) \) a.s. with a measurable function of \((w(r))_{0 \leq r \leq t}\);
2. for every \( w \in C(\mathbb{R}^+; \mathbb{R}^m) \), the mapping \( x \mapsto F(x, w) \) is continuous;
3. for every \( x \in \mathbb{R}^d \), for every choice of the (complete) filtered probability space \((\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t \geq 0})\) and of the \( m \times \) dimensional \((\mathcal{F}_t)_{t \geq 0}\)-Brownian motion \( B \) with \( B_0 = 0 \), the process \( X_t \) defined \( X_t = F(x, B)_t \) is the unique solution of \( E_x(\sigma, b) \); furthermore, if \( U \) is an \( \mathcal{F}_0 \)-measurable real random variable, the process \( F(U, B)_t \) is the unique solution with \( X_0 = U \).

For \( x \in \mathbb{R}^d \), let \( Y^x \) be a solution of \( E_x(\sigma, b) \). Weak uniqueness implies that for every \( t \geq 0 \), the law of \( Y^x_t \) does not depend on the choice of the solution. Moreover, this law is the image of Wiener measure on \( C(\mathbb{R}^+; \mathbb{R}^d) \) under the mapping \( w \mapsto F(x, w)_t \). We can easily check the following result.

**Theorem 25.2.** Assume that \( X = (X_t)_{t \geq 0} \) is a solution of \( E(\sigma, b) \) on a complete filtered probability space \((\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t \geq 0})\). Then for any \( t \geq 0 \) we have

\[
(X_{t+s})_{t \geq 0} \mid \mathcal{F}_s \overset{\text{d}}{=} Y^x_s
\]

where \( Y^x \) is any solution of \( E_x(\sigma, b) \) with \( x \in \mathbb{R} \) fixed.

**Proof.** For fixed \( s \geq 0 \), we have for all \( t \geq 0 \)

\[
X_{t+s} = X_s + \int_s^{t+s} (X_r)dr + \int_s^{t+s} \sigma(X_r)dB_r
\]

where \( B \) is an \((\mathcal{F}_t)_{t \geq 0}\)-Brownian motion starting from 0. We define, for every \( t \geq 0 \)

\[
X'_t = X_{t+s}, \quad \mathcal{F}'_t = \mathcal{F}_{t+s}, \quad B'_t = B_{t+s} - B_s.
\]

We note that the filtration \((\mathcal{F}'_t)_{t \geq 0}\) is complete, the process \( X' \) is adapted to \((\mathcal{F}'_t)_{t \geq 0}\) and \( B' \) is a \((\mathcal{F}'_t)_{t \geq 0}\)-Brownian motion starting from 0. Furthermore, using results for the stochastic integral of adapted processes with continuous sample paths, we get

\[
\int_s^{t+s} \sigma(X_r)dB_r = \int_0^t \sigma(X'_r)dB'_r
\]

where the stochastic integral in the rhs is computed in the new filtration. Thus, we have

\[
X'_t = X_s + \int_0^t (X'_r)dr + \int_0^t \sigma(X'_r)dB'_r.
\]
Hence $X'$ solves $E(\sigma, b)$ on the space $(\Omega, \mathcal{F}, P, (\mathcal{F}_t')_{t \geq 0})$ with the Brownian motion $B'$ and initial value $X'_0 = X_s$, which is $\mathcal{F}_0'$-measurable. By the previous theorem, we get $X' = F(X_s, B')$ a.s. ■

Observe that in the above proof $B'$ is independent of $\mathcal{F}_s$ and thus given the complete history up to time $s$, the distribution of the future of the process $X$ depends only on $X_s$. This is the starting point for the definition of Markov processes, which we will study now. We only study time-homogeneous Markov processes here.

### 25.1 Markov processes

Markov processes form a fundamental class of stochastic processes, with many applications in real life problems. The reason why Markov processes are so important comes from the so-called Markov property, the future depends on the past only through the present. This enables many explicit calculations that would be intractable for more general stochastic processes. Heuristically, for a Markov process $X$, for any $s, t \geq 0$ and event $A$, $P(X_{t+s} \in A \mid \mathcal{F}_s)$ should be a $[0,1]$-valued measurable function of $X_s$. Also, this clearly depends on $A$ and is a probability measure as a function of $A$. We formally define these functions as transition kernels.

**Definition 25.3 (Markov transition kernel).** Let $(E, \mathcal{E})$ be a measurable space. A *Markov transition kernel*, or simply *transition kernel*, from $E$ into $E$ is a mapping $Q : E \times E \to [0,1]$ satisfying the following two properties:

1. For every $x \in E$, the mapping $A \mapsto Q(x, A)$ is a probability measure on $(E, \mathcal{E})$.

2. For every $A \in \mathcal{E}$, the mapping $x \mapsto Q(x, A)$ is $\mathcal{E}$-measurable.

Note that, when $E$ is countable with $\mathcal{E} = 2^E$, $Q$ is characterized by the transition matrix $(Q(x, \{y\}))_{x,y \in E})$. Moreover, for any bounded measurable real-valued function $f$ on $E$, we can define the bounded measurable function

$$Qf(x) := \int Q(x, dy) f(y), \quad x \in E.$$ 

Moreover, we have $\|Qf\|_\infty \leq \|f\|_\infty$. Thus $Q$ can be considered as a contractive operator on $B(E)$, the space of bounded measurable functions on $E$ equipped with the supremum norm. Note that, $Q$ maps positive functions to positive functions.  

**Definition 25.4 (transition semigroup).** A collection $(Q_t)_{t \geq 0}$ of transition kernels on $E$ is called a *transition semigroup* if the following three properties hold.

1. For every $x \in E$, $Q_0(x, dy) = \delta_x(dy)$ or $Q_0$ is the identity operator.

2. For every $s, t \geq 0$ and $A \in \mathcal{E}$,

$$Q_{t+s}(x, A) = \int Q_t(x, dy) Q_s(y, A) \quad \text{or} \quad Q_{t+s} = Q_t \circ Q_s.$$ 

3. For every $A \in \mathcal{E}$, the function $(t, x) \mapsto Q_t(x, A)$ is measurable with respect to the $\sigma$-field $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{E}$. 

Condition 2. is known as the Chapman–Kolmogorov identity. Also, one can view a transition semigroup as semigroup of contractions of \( \mathcal{B}(E) \). Now, we are ready to define a Markov Process.

**Definition 25.5 (Markov Process).** Let \((Q_t)_{t \geq 0}\) be a transition semigroup on \(E\). A Markov process with respect to the filtration \((\mathcal{F}_t)_{t \geq 0}\) and with transition semigroup \((Q_t)_{t \geq 0}\) is an \((\mathcal{F}_t)\)-adapted process \((X_t)_{t \geq 0}\) with values in \(E\) such that, for every \(s, t \geq 0\) and \(f \in \mathcal{B}(E)\),

\[
\mathbb{E}(f(X_{t+s}) \mid \mathcal{F}_s) = Q_tf(X_s) \text{ a.s.}
\]

In particular, we have

\[
\mathbb{P}(X_{t+s} \in A \mid \mathcal{F}_s) = Q_t(X_s, A).
\]

In general, we will take the canonical filtration as \((\mathcal{F}_t)_{t \geq 0}\) unless explicitly mentioned. As a consequence, we can explicitly write down the joint distribution of \(X\) at finitely many time points and using Kolmogorov extension theorem one can prove the existence of a Markov process with transition semigroup \((Q_t)_{t \geq 0}\).

As an example, with \(E = \mathbb{R}^d, \mathcal{E} = \mathcal{B}(\mathbb{R}^d)\), the standard Brownian motion \(B\) is a Markov process with

\[
Q_tf(x) = \mathbb{E}(f(B_{t+s}) \mid B_s = x) = \mathbb{E}f(x + \sqrt{t}\eta)
\]

where \(\eta \sim N_d(0, 1)\).

Note that, for a continuous function \(g : [0, \infty) \to [0, 1]\) satisfying \(g(t+s) = g(t)g(s)\) for all \(s, t \geq 0\) and \(g(0) = 1\), \(g\) must be of the form \(g(t) = e^{\alpha t}\) for some real number \(\alpha \leq 0\). The number \(\alpha\) completely describes the function \(g\). Moreover, \(\alpha\) can be recovered from \(g\) by looking at \(g(0)\) (which is not obvious for operators) or by looking at the resolvent or the Laplace transform (linear operation) \(R_\lambda g = \int_0^\infty e^{-\lambda t}g(t)dt = (\lambda - \alpha)^{-1}\) for \(\lambda > 0\) and noting that \(\alpha = \lambda - R_\lambda^{-1}\) does not depend on \(\lambda > 0\). We can generalize the above approaches to the transition semigroup under appropriate assumptions.

First, we introduce the important notion of the resolvent.

**Definition 25.6 (Resolvent).** For \(\lambda > 0\), the \(\lambda\)-Resolvent of the transition semigroup \((Q_t)_{t \geq 0}\) is the linear operator \(R_\lambda : \mathcal{B}(E) \to \mathcal{B}(E)\) defined as

\[
R_\lambda f(x) := \int_0^\infty e^{-\lambda t}Q_tf(x)\,dx, \quad f \in \mathcal{B}(E), x \in E.
\]

Property 3. of the definition of a transition semigroup is used here to get the measurability of the mapping \(t \mapsto Q_t f(x)\) and define \(R_\lambda\) rigorously. We have the following.

**Lemma 25.7.**

1. \(||\lambda R_\lambda f|| \leq ||f||.\)

2. (Positivity.) If \(0 \leq f \leq 1\), then \(0 \leq R_\lambda f \leq 1\).

3. (Resolvent equation.) If \(\lambda, \mu > 0\), we have \(R_\lambda - R_\mu = -(\lambda - \mu)R_\lambda R_\mu\).

**Proof.** The first two properties are obvious as \(\lambda R_\lambda f(x) = \mathbb{E}Q_\tau f(x)\) where \(\tau\) is an independent
Exponential rate r.v. To prove 3., take $\lambda > \mu$ and note that by the semigroup property we get

$$R_\lambda R_\mu f(x) = \int_0^\infty e^{-\lambda s} Q_s \left( \int_0^\infty e^{-\mu t} Q_t f(t) \, dt \right) (x) ds$$

$$= \int_0^\infty \int_0^\infty e^{-\lambda s} e^{-\mu t} Q_{s+t} f(x) \, dt ds$$

$$= \int_0^\infty \int_s^\infty e^{-\lambda s} e^{-\mu(u-s)} Q_u f(x) \, duds$$

$$= \int_0^\infty e^{-\mu u} Q_u f(x) \int_0^u e^{-(\lambda-\mu)s} \, ds \, du.$$

Now using $-(\lambda - \mu) e^{-\mu u} \int_0^u e^{-(\lambda-\mu)s} \, ds = e^{-\lambda u} - e^{-\mu u}$ and the definition of Resolvent, we get the result.

As a consequence, we get that $R_\lambda$'s commute with each other. Note that, for Standard Brownian Morion we get with $\tau \sim \text{Exponential}(1)$, $\eta \sim N(0, 1)$, independent,

$$R_\lambda f(x) = \lambda^{-1} E Q_{\tau/\lambda} f(x) = \lambda^{-1} E f(x + \sqrt{\tau/\lambda} \cdot \eta) = \lambda^{-1} E f(x + \lambda^{-1/2} \cdot \xi)$$

where $\xi \overset{d}{=} \sqrt{\tau} \cdot \eta$ has Laplace$(0, 1/\sqrt{2})$ distribution (follows easily from characteristic function computation).