Math 562 Fall 2020

Renming Song

University of Illinois at Urbana-Champaign

October 28, 2020

Outline

Outline

HW5 is due Friday Nov. 6 at noon. Please submit your homework via the course Moodle page. Make sure that your HW is uploaded successfully.

Outline

The goal of this chapter is to give a concise introduction to the main ideas of the theory of continuous time Markov processes. Markov processes form a fundamental class of stochastic processes, with many applications in real life problems outside mathematics. The reason why Markov processes are so important comes from the so-called Markov property, which enables many explicit calculations that would be intractable for more general random processes. Although the theory of Markov processes is by no means the central topic of this book, it will play a significant role in the next chapters, in particular in our discussion of stochastic differential equations. In fact the whole invention of Ito's stochastic calculus was motivated by the study of the Markov processes obtained as solutions of stochastic differential equations, which are also called diffusion processes.

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

(i) For each $x \in E$, the mapping $\mathcal{E} \ni A \mapsto Q(x, A) \in [0, 1]$ is a probability measure on $(E, \mathcal{E})$.

(ii) For every $A \in \mathcal{E}$, $E \ni x \mapsto Q(x, A) \in [0, 1]$ is $\mathcal{E}$-measurable.

In what follows we say transition kernel instead of Markov transition kernel.

In the case where E is finite or countable (and equipped with the σ-field of all subsets of E), Q is characterized by the matrix $(Q(x, \{y\})_{x,y \in E}$.

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

(i) For each $x \in E$, the mapping $\mathcal{E} \ni A \mapsto Q(x, A) \in [0, 1]$ is a probability measure on $(E, \mathcal{E})$.

(ii) For every $A \in \mathcal{E}$, $E \ni x \mapsto Q(x, A) \in [0, 1]$ is $\mathcal{E}$-measurable.

In what follows we say transition kernel instead of Markov transition kernel.

In the case where E is finite or countable (and equipped with the σ-field of all subsets of E), Q is characterized by the matrix $(Q(x, \{y\})_{x,y \in E}$.

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

(i) For each $x \in E$, the mapping $\mathcal{E} \ni A \mapsto Q(x, A) \in [0, 1]$ is a probability measure on $(E, \mathcal{E})$.

(ii) For every $A \in \mathcal{E}$, $E \ni x \mapsto Q(x, A) \in [0, 1]$ is $\mathcal{E}$-measurable.

In what follows we say transition kernel instead of Markov transition kernel.

In the case where E is finite or countable (and equipped with the σ-field of all subsets of E), Q is characterized by the matrix $(Q(x, \{y\})_{x,y \in E}$.

If $f : E \to \mathbb{R}$ is bounded and measurable (resp. non-negative and measurable), the function Qf defined by

$$Qf(x) = \int Q(x, dy) f(y)$$

is also bounded and measurable (resp. non-negative and measurable) on E. Indeed, if f is an indicator function, the measurability of Qf is just property (ii) and the general case follows from standard arguments.

Definition 6.1

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

(i) For every $x \in E$, $Q_0(x, dy) = \delta_x(dy)$.

(ii) For every $s, t \geq 0$ and $A \in \mathcal{E}$,

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

The equation above is called the Chapman-Kolmogorov equation.

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

Let $B(E)$ be the vector space of all bounded measurable real-valued functions on E, equipped with the norm $\|f\| = \sup\{|f(x)| : x \in E\}$. Then the linear mapping $B(E) \ni f \mapsto Q_t f$ is a contraction of $B(E)$. From this point of view, the Chapman-Kolmogorov equation is equivalent to the relation

$$Q_{t+s} = Q_t Q_s$$

for every $s, t \geq 0$. This allows one to view $(Q_t)_{t\geq 0}$ as a semigroup of contractions of $B(E)$.

We now consider a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,\infty]}, \mathbb{P})$.

Let $B(E)$ be the vector space of all bounded measurable real-valued functions on E, equipped with the norm $\|f\| = \sup\{|f(x)| : x \in E\}$. Then the linear mapping $B(E) \ni f \mapsto Q_t f$ is a contraction of $B(E)$. From this point of view, the Chapman-Kolmogorov equation is equivalent to the relation

$$Q_{t+s} = Q_t Q_s$$

for every $s, t \geq 0$. This allows one to view $(Q_t)_{t \geq 0}$ as a semigroup of contractions of $B(E)$.

We now consider a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0,\infty]}, \mathbb{P})$.

Definition 6.2

Let $(Q_t)_{t\geq 0}$ be a transition semigroup on E. A Markov process (with respect to the filtration $(\mathcal{F}_t)$) 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 $f \in B(E)$,

$$\mathbb{E}[f(X_{s+t})|\mathcal{F}_s] = Q_t f(X_s).$$

When we speak about a Markov process X without specifying the filtration, we implicitly mean that the property of the definition holds with the canonical filtration $\mathcal{F}_t^X = \sigma\{X_r : 0 \leq r \leq t\}$. We note that, if X is a Markov process with respect to a filtration $(\mathcal{F}_t)$, it is automatically also a Markov process (with the same semigroup) with respect to $(\mathcal{F}_t^X)$.

Definition 6.2

Let $(Q_t)_{t\geq 0}$ be a transition semigroup on E. A Markov process (with respect to the filtration $(\mathcal{F}_t)$) 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 $f \in B(E)$,

$$\mathbb{E}[f(X_{s+t})|\mathcal{F}_s] = Q_t f(X_s).$$

When we speak about a Markov process X without specifying the filtration, we implicitly mean that the property of the definition holds with the canonical filtration $\mathcal{F}_t^X = \sigma\{X_r : 0 \leq r \leq t\}$. We note that, if X is a Markov process with respect to a filtration $(\mathcal{F}_t)$, it is automatically also a Markov process (with the same semigroup) with respect to $(\mathcal{F}_t^X)$.

The definition of a Markov process can be interpreted as follows. Taking $f = 1_A$, we have

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

and in particular

$$\mathbb{P}(X_{s+t} \in A|X_r : 0 \leq r \leq s] = Q_t(X_s, A).$$

Hence the conditional distribution of X_{s+t} knowing the past $(X_r : 0 \leq r \leq s)$ before time s is given by $Q_t(X_s, \cdot)$, and this conditional distribution only depends on the present state X_s. This is the Markov property (informally, if one wants to predict the future after time s, the past up to time s does not give more information than just the present at time s).

Consequences of the definition

Let $\gamma(dx)$ be the law of X_0. If $0 < t_1 < t_2 < \cdots < t_p$ and $A_0, A_1, \ldots, A_p \in \mathcal{E}$,

$$\mathbb{P}(X_0 \in A_0, X_{t_1} \in A_1, \ldots, X_{t_p} \in A_p)$$

$$= \int_{A_0} \gamma(dx_0) \int_{A_1} Q_{t_1}(x_0, dx_1) \int_{A_2} Q_{t_2 - t_1}(x_2, dx_2) \cdots \int_{A_p} Q_{t_p - t_{p-1}}(x_{p-1}, dx_p).$$

More generally, if $f_0, f_1, \ldots, f_p \in B(E)$,

$$\mathbb{E}[f_0(X_0) f_1(X_{t_1}) \cdots f_p(X_{t_p})] = \int \gamma(dx_0) f_0(x_0) \int Q_{t_1}(x_0, dx_1) f_1(x_1)$$

$$\int Q_{t_2 - t_1}(x_2, dx_2) f_2(x_2) \cdots \int Q_{t_p - t_{p-1}}(x_{p-1}, dx_p) f_p(x_p).$$

Consequences of the definition

Let $\gamma(dx)$ be the law of X_0. If $0 < t_1 < t_2 < \cdots < t_p$ and $A_0, A_1, \ldots, A_p \in \mathcal{E}$,

$$\mathbb{P}(X_0 \in A_0, X_{t_1} \in A_1, \ldots, X_{t_p} \in A_p)$$
$$= \int_{A_0} \gamma(dx_0) \int_{A_1} Q_{t_1}(x_0, dx_1) \int_{A_2} Q_{t_2-t_1}(x_2, dx_2) \cdots \int_{A_p} Q_{t_p-t_{p-1}}(x_{p-1}, dx_p).$$

More generally, if $f_0, f_1, \ldots, f_p \in B(E)$,

$$\mathbb{E}[f_0(X_0)f_1(X_{t_1}) \cdots f_p(X_{t_p})] = \int \gamma(dx_0) f_0(x_0) \int Q_{t_1}(x_0, dx_1) f_1(x_1)$$
$$\int Q_{t_2-t_1}(x_2, dx_2) f_2(x_2) \cdots \int Q_{t_p-t_{p-1}}(x_{p-1}, dx_p) f_p(x_p).$$

The last formula is derived from the definition by induction on p. Note that, conversely, if the latter formula holds for any choice of $0 < t_1 < t_2 < \cdots < t_p$ and $f_0, f_1, \ldots, f_p \in B(E)$, then $(X_t)_{t\geq 0}$ is a Markov process of semigroup $(Q_t)_{t\geq 0}$, with respect to its canonical filtration $(\mathcal{F}_t^X)$.

From the preceding formulas, we see that the finite-dimensional marginals of the process X are completely determined by the semigroup $(Q_t)_{t\geq 0}$ and the law of X_0 (initial distribution).

The last formula is derived from the definition by induction on p. Note that, conversely, if the latter formula holds for any choice of $0 < t_1 < t_2 < \cdots < t_p$ and $f_0, f_1, \ldots, f_p \in B(E)$, then $(X_t)_{t\geq 0}$ is a Markov process of semigroup $(Q_t)_{t\geq 0}$, with respect to its canonical filtration $(\mathcal{F}_t^X)$.

From the preceding formulas, we see that the finite-dimensional marginals of the process X are completely determined by the semigroup $(Q_t)_{t\geq 0}$ and the law of X_0 (initial distribution).

Example

If $E = \mathbb{R}^d$, we can take, for every $t > 0$ and $x \in \mathbb{R}^d$,

$$Q_t(x, dy) = p_t(y - x)dy$$

where, for $z \in \mathbb{R}^d$,

$$p_t(z) = (2\pi t)^{-d/2} \exp\left(-\frac{|z|^2}{2}\right).$$

It is straightforward to verify that this defines a transition semigroup on $\mathbb{R}^d$ and the associated Markov process is a d-dim pre-Brownian motion.

We now address the problem of the existence of a Markov process with a given semigroup. To this end, we will need a general theorem of construction of random processes, namely the Kolmogorov extension theorem. We give without proof the special case of this theorem that is of interest to us.

Example

If $E = \mathbb{R}^d$, we can take, for every $t > 0$ and $x \in \mathbb{R}^d$,

$$Q_t(x, dy) = p_t(y - x)dy$$

where, for $z \in \mathbb{R}^d$,

$$p_t(z) = (2\pi t)^{-d/2} \exp\left(-\frac{|z|^2}{2}\right).$$

It is straightforward to verify that this defines a transition semigroup on $\mathbb{R}^d$ and the associated Markov process is a d-dim pre-Brownian motion.

We now address the problem of the existence of a Markov process with a given semigroup. To this end, we will need a general theorem of construction of random processes, namely the Kolmogorov extension theorem. We give without proof the special case of this theorem that is of interest to us.

Let $\Omega^* = E^{\mathbb{R}_+}$ be the space of all mappings $\omega : \mathbb{R}_+ \to E$. We equip Ω^* with the σ-field $\mathcal{F}^*$ generated by the coordinate mappings $\omega \mapsto \omega(t)$ for $t \in \mathbb{R}_+$. Let $F(\mathbb{R}_+)$ be the collection of all finite subsets of $\mathbb{R}_+$, and, for every $U \in F(\mathbb{R}_+)$, let $\pi_U : \Omega^* \to E^U$ be the mapping which associates with every $\omega : \mathbb{R}_+ \to E$ its restriction to U. If $U; V \in F(\mathbb{R}_+)$ and $U \subset V$, we similarly write $\pi_U^V : E^V \to E^U$ for the obvious restriction mapping.

Recall that a topological space is Polish if its topology is separable (there exists a dense sequence) and can be defined by a complete metric.

Let $\Omega^* = E^{\mathbb{R}_+}$ be the space of all mappings $\omega : \mathbb{R}_+ \to E$. We equip Ω^* with the σ-field $\mathcal{F}^*$ generated by the coordinate mappings $\omega \mapsto \omega(t)$ for $t \in \mathbb{R}_+$. Let $F(\mathbb{R}_+)$ be the collection of all finite subsets of $\mathbb{R}_+$, and, for every $U \in F(\mathbb{R}_+)$, let $\pi_U : \Omega^* \to E^U$ be the mapping which associates with every $\omega : \mathbb{R}_+ \to E$ its restriction to U. If $U; V \in F(\mathbb{R}_+)$ and $U \subset V$, we similarly write $\pi_U^V : E^V \to E^U$ for the obvious restriction mapping.

Recall that a topological space is Polish if its topology is separable (there exists a dense sequence) and can be defined by a complete metric.

Theorem 6.3

Assume that E is a Polish space equipped with its Borel σ-field $\mathcal{E}$. For every $U \in F(\mathbb{R}_+)$, let μ_U be a probability measure on E^U. Assume that the collection $(\mu_U : U \in F(\mathbb{R}_+))$ is consistent in the following sense: If $U \subset V$, μ_U is the image of μ_V under the mapping π_U^V. Then there exists a unique probability measure μ on $(\Omega^*, \mathcal{F}^*)$ such that $\pi_U(\mu) = \mu_U$ for every $U \in F(\mathbb{R}_+)$.

The Kolmogorov extension theorem allows one to construct stochastic processes having prescribed finite-dimensional marginals.

Theorem 6.3

Assume that E is a Polish space equipped with its Borel σ-field $\mathcal{E}$. For every $U \in F(\mathbb{R}_+)$, let μ_U be a probability measure on E^U. Assume that the collection $(\mu_U : U \in F(\mathbb{R}_+))$ is consistent in the following sense: If $U \subset V$, μ_U is the image of μ_V under the mapping π_U^V. Then there exists a unique probability measure μ on $(\Omega^*, \mathcal{F}^*)$ such that $\pi_U(\mu) = \mu_U$ for every $U \in F(\mathbb{R}_+)$.

The Kolmogorov extension theorem allows one to construct stochastic processes having prescribed finite-dimensional marginals.

Let $(X_t)_{t \geq 0}$ be the canonical process on Ω^*:

$$X_t(\omega) = \omega(t), \qquad t \geq 0.$$

If μ is a probability measure on Ω^* and $U = \{t_1, \ldots, t_p\} \in F(\mathbb{R}_+)$, with $t_1 < \cdots < t_p$, then $(X_{t_1}, \ldots, X_{t_p})$ can be viewed as a random variable with values in E^U, provided we identify E^U with E^p via the mapping $\omega \to (\omega(t_1), \ldots, \omega(t_p))$. Furthermore, the distribution of $(X_{t_1}, \ldots, X_{t_p})$ under μ is $\pi_U(\mu)$. The Kolmogorov theorem can thus be rephrased by saying that given a collection $(\mu_U : U \in F(\mathbb{R}_+))$ of finite-dimensional marginal distributions, which satisfies the consistency condition (this condition is clearly necessary for the desired conclusion), one can construct a probability measure μ on Ω^*, under which the finite-dimensional marginals of the canonical process X are the measures μ_U.

Corollary 6.4

We assume that E satisfies the assumption of the previous theorem and that $(Q_t)_{t\geq 0}$ is a transition semigroup on E. Let γ be a probability measure on E. Then there exists a (unique) probability measure $\mathbb{P}$ on Ω^* under which the canonical process $(X_t)_{t\geq 0}$ is a Markov process with transition semigroup $(Q_t)_{t\geq 0}$ and initial law γ.

Proof of Corollary 6.4

Let $U = \{t_1, \ldots, t_p\} \in F(\mathbb{R}_+)$, with $t_1 < \cdots < t_p$. We define a probability measure $\mathbb{P}_U$ on E^U (identified with E^p as explained above) by setting for every measurable subset A of E^U:

$$\int P_U(dx_1, \ldots, dx_p) 1_A(x_1, \ldots, x_p)$$

$$= \int \gamma(dx_0) \int Q_{t_1}(x_0, dx_1) \int Q_{t_2 - t_1}(x_1, dx_2) \cdots \int Q_{t_p - t_{p-1}}(x_{p-1}, dx_p).$$

Corollary 6.4

We assume that E satisfies the assumption of the previous theorem and that $(Q_t)_{t\geq 0}$ is a transition semigroup on E. Let γ be a probability measure on E. Then there exists a (unique) probability measure $\mathbb{P}$ on Ω^* under which the canonical process $(X_t)_{t\geq 0}$ is a Markov process with transition semigroup $(Q_t)_{t\geq 0}$ and initial law γ.

Proof of Corollary 6.4

Let $U = \{t_1, \ldots, t_p\} \in F(\mathbb{R}_+)$, with $t_1 < \cdots < t_p$. We define a probability measure $\mathbb{P}_U$ on E^U (identified with E^p as explained above) by setting for every measurable subset A of E^U:

$$\int P_U(dx_1, \ldots, dx_p) 1_A(x_1, \ldots, x_p)$$
$$= \int \gamma(dx_0) \int Q_{t_1}(x_0, dx_1) \int Q_{t_2 - t_1}(x_1, dx_2) \cdots \int Q_{t_p - t_{p-1}}(x_{p-1}, dx_p).$$

Proof of Corollary 6.4 (cont)

Using the Chapman-Kolmogorov relation, one verifies that the measures $\mathbb{P}_U$ satisfy the consistency condition. The Kolmogorov theorem then gives the existence (and uniqueness) of a probability measure $\mathbb{P}$ on Ω^* whose finite-dimensional marginals are the measures $\mathbb{P}_U$, $U \in F(\mathbb{R}_+)$. By a previous observation, this implies that $(X_t)_{t\geq 0}$ is under $\mathbb{P}$ a Markov process with semigroup $(Q_t)_{t\geq 0}$ with respect to the canonical filtration.

For $x \in E$, let $\mathbb{P}_x$ be the measure given by the preceding corollary when $\gamma = \delta_x$. Then, the mapping $x \mapsto \mathbb{P}_x$ is measurable in the sense that $x \mapsto \mathbb{P}_x(A)$ is measurable, for every $A \in \mathcal{F}^*$. In fact, the latter property holds when A only depends on a finite number of coordinates (in that case, there is an explicit formula for $\mathbb{P}_x(A)$) and a monotone class argument gives the general case.

Proof of Corollary 6.4 (cont)

Using the Chapman-Kolmogorov relation, one verifies that the measures $\mathbb{P}_U$ satisfy the consistency condition. The Kolmogorov theorem then gives the existence (and uniqueness) of a probability measure $\mathbb{P}$ on Ω^* whose finite-dimensional marginals are the measures $\mathbb{P}_U$, $U \in F(\mathbb{R}_+)$. By a previous observation, this implies that $(X_t)_{t\geq 0}$ is under $\mathbb{P}$ a Markov process with semigroup $(Q_t)_{t\geq 0}$ with respect to the canonical filtration.

For $x \in E$, let $\mathbb{P}_x$ be the measure given by the preceding corollary when $\gamma = \delta_x$. Then, the mapping $x \mapsto \mathbb{P}_x$ is measurable in the sense that $x \mapsto \mathbb{P}_x(A)$ is measurable, for every $A \in \mathcal{F}^*$. In fact, the latter property holds when A only depends on a finite number of coordinates (in that case, there is an explicit formula for $\mathbb{P}_x(A)$) and a monotone class argument gives the general case.

Moreover, for any probability measure γ on E, the measure defined by

$$\mathbb{P}_{(\gamma)}(A) = \int \gamma(dx)\mathbb{P}_x(A)$$

is the unique probability measure on Ω^* under which the canonical process $(X_t)_{t\geq 0}$ is a Markov process with semigroup $(Q_t)_{t\geq 0}$ and initial law γ.

Summarizing, the preceding corollary allows one to construct (under a topological assumption on E) a Markov process $(X_t)_{t\geq 0}$ with semigroup $(Q_t)_{t\geq 0}$, which starts with a given initial distribution. More precisely, we get a measurable collection of probability measures $(\mathbb{P}_x)_{x\in E}$ such that the Markov process X starts from x under $\mathbb{P}_x$. However, a drawback of the method that we used is the fact that it does not give any information on the regularity properties of sample paths of X, we cannot even assert that these sample paths are measurable.

Moreover, for any probability measure γ on E, the measure defined by

$$\mathbb{P}_{(\gamma)}(A) = \int \gamma(dx)\mathbb{P}_x(A)$$

is the unique probability measure on Ω^* under which the canonical process $(X_t)_{t\geq 0}$ is a Markov process with semigroup $(Q_t)_{t\geq 0}$ and initial law γ.

Summarizing, the preceding corollary allows one to construct (under a topological assumption on E) a Markov process $(X_t)_{t\geq 0}$ with semigroup $(Q_t)_{t\geq 0}$, which starts with a given initial distribution. More precisely, we get a measurable collection of probability measures $(\mathbb{P}_x)_{x\in E}$ such that the Markov process X starts from x under $\mathbb{P}_x$. However, a drawback of the method that we used is the fact that it does not give any information on the regularity properties of sample paths of X, we cannot even assert that these sample paths are measurable.