Math 562 Fall 2020

Renming Song

University of Illinois at Urbana-Champaign

November 30, 2020

Outline

Outline

In this section, we consider the homogeneous case where $\sigma(t, y) = \sigma(y)$ and $b(t, y) = b(y)$. We assume that σ and b are Lipschitz: There exists a constant K such that, for every $x, y \in \mathbb{R}^d$,

$$|\sigma(x) - \sigma(y)| \leq K|x - y|, \quad |b(x) - b(y)| \leq K|x - y|.$$

Let $x \mapsto F_x$ be the mappings defined in Theorem 8.5. Let X be a solution of $E(\sigma, b)$ on a (complete) filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$. We know that $(X_t)_{t\geq 0}$ is a Markov process with respect to the filtration $(\mathcal{F}_t)$, with semigroup $(Q_t)_{t\geq 0}$ defined by

$$Q_t f(x) = \mathbb{E}[f(X_t^x)],$$

where X^x is an arbitrary solution of $E_x(\sigma, b)$. Q_t is also given by

$$Q_t f(x) = \int f(F_x(w)_t) W(dw). \tag{1}$$

In this section, we consider the homogeneous case where $\sigma(t, y) = \sigma(y)$ and $b(t, y) = b(y)$. We assume that σ and b are Lipschitz: There exists a constant K such that, for every $x, y \in \mathbb{R}^d$,

$$|\sigma(x) - \sigma(y)| \leq K|x - y|, \quad |b(x) - b(y)| \leq K|x - y|.$$

Let $x \mapsto F_x$ be the mappings defined in Theorem 8.5. Let X be a solution of $E(\sigma, b)$ on a (complete) filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$. We know that $(X_t)_{t\geq 0}$ is a Markov process with respect to the filtration $(\mathcal{F}_t)$, with semigroup $(Q_t)_{t\geq 0}$ defined by

$$Q_t f(x) = \mathbb{E}[f(X_t^x)],$$

where X^x is an arbitrary solution of $E_x(\sigma, b)$. Q_t is also given by

$$Q_t f(x) = \int f(F_x(w)_t) W(dw). \tag{1}$$

We use $C_c^2(\mathbb{R}^d)$ to denote the space of all twice continuously differentiable functions with compact support on $\mathbb{R}^d$.

Theorem 8.7

The semigroup $(Q_t)_{t\geq 0}$ is Feller. Furthermore, its generator L is such that

$$C_c^2(\mathbb{R}^d) \subset D(L)$$

and, for every $f \in C_c^2(\mathbb{R}^d)$,

$$Lf(x) = \frac{1}{2}\sum_{i,j=1}^{d}(\sigma\sigma^*)_{ij}(x)\frac{\partial^2 f}{\partial x_i \partial x_j}(x) + \sum_{i=1}^{d} b_i(x)\frac{\partial f}{\partial x_i}(x).$$

We use $C_c^2(\mathbb{R}^d)$ to denote the space of all twice continuously differentiable functions with compact support on $\mathbb{R}^d$.

Theorem 8.7

The semigroup $(Q_t)_{t\geq 0}$ is Feller. Furthermore, its generator L is such that

$$C_c^2(\mathbb{R}^d) \subset D(L)$$

and, for every $f \in C_c^2(\mathbb{R}^d)$,

$$Lf(x) = \frac{1}{2}\sum_{i,j=1}^{d}(\sigma\sigma^*)_{ij}(x)\frac{\partial^2 f}{\partial x_i \partial x_j}(x) + \sum_{i=1}^{d} b_i(x)\frac{\partial f}{\partial x_i}(x).$$

Proof of Theorem 8.7

We give the proof only in the case when σ and b are bounded. The general case can be dealt with using a localization argument. We fix $f \in C_0(\mathbb{R}^d)$ and we first show that $Q_t f \in C_0(\mathbb{R}^d)$. Since the mappings $x \mapsto F_x$ are continuous, formula (1) and dominated convergence show that $Q_t f$ is continuous. Then, since

$$X_t^x = x + \int_0^t \sigma(X_s^x) dB_s + \int_0^t b(X_s^x) ds,$$

and σ and b are bounded, we get the existence of a constant C, which does not depend on t; x, such that

$$\mathbb{E}[(X_t^x - x)^2] \leq C(t + t^2). \tag{2}$$

Using Markov's inequality, we have thus, for every $t \geq 0$, as $A \to \infty$,

$$\sup_{x \in \mathbb{R}^d} \mathbb{P}(|X_t^x - x| > A) \to 0.$$

Proof of Theorem 8.7 (cont)

Since

$$|Q_t f(x)| = |\mathbb{E}[f(X_t^x)]| \leq |\mathbb{E}[f(X_t^x) 1_{\{|X_t^x - x| \leq A\}}]| + \|f\| \mathbb{P}(|X_t^x - x| > A),$$

we get, using our assumption $f \in C_0(\mathbb{R}^d)$,

$$\limsup_{x \to \infty} |Q_t f(x)| \leq \|f\| \sup_{x \in \mathbb{R}^d} \mathbb{P}(|X_t^x - x| > A),$$

and thus, since A was arbitrary,

$$\lim_{x \to \infty} Q_t f(x) = 0,$$

which completes the proof of the property $Q_t f \in C_0(\mathbb{R}^d)$.

Proof of Theorem 8.7 (cont)

Let us show similarly that $Q_t f \to f$ as $t \to 0$. For every $\epsilon > 0$,

$$\sup_{x\in\mathbb{R}^d} |\mathbb{E}[f(X_t^x)] - f(x)| \le \sup_{x,y\in\mathbb{R}^d, ||x-y||\le\epsilon} |f(y)-f(x)| + 2\|f\| \sup_{x\in\mathbb{R}^d} \mathbb{P}(|X_t^x - x| > \epsilon).$$

However, using (2) and Markov's inequality again, we get that as $t \to 0$,

$$\sup_{x\in\mathbb{R}^d} \mathbb{P}(|X_t^x - x| > \epsilon) \to 0,$$

hence

$$\limsup_{t\to 0} \|Q_t f - f\| = \limsup_{t\to 0} \left(\sup_{x\in\mathbb{R}^d} |\mathbb{E}[f(X_t^x)] - f(x)| \right)$$
$$\le \sup_{x,y\in\mathbb{R}^d, ||x-y||\le\epsilon} |f(y) - f(x)|$$

which can be made arbitrarily close to 0 by taking ϵ small.

Proof of Theorem 8.7 (cont)

Let us prove the second assertion of the theorem. Let $f \in C_c^2(\mathbb{R}^d)$. We apply Ito's formula to $f(X_t^x)$, recalling that, if $X_t^x = (X_t^{x,1}, \dots, X_t^{x,d})$, we have, for every $i \in \{1, \dots, d\}$,

$$X_t^{x,i} = x_i + \sum_{j=1}^m \int_0^t \sigma_{ij}(X_s^x) dB_s^j + \int_0^t b_i(X_s^x) ds.$$

We get

$$f(X_t^x) = f(x) + M_t + \sum_{i=1}^d \int_0^t b_i(X_s^x) \frac{\partial f}{\partial x_i}(X_s^x) ds$$
$$+ \frac{1}{2} \sum_{i,k=1}^d \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_k}(X_s^x) d\langle X^{x,i}, X^{x,k} \rangle_s$$

where M is a continuous local martingale. Moreover, if $i, k \in \{1, \dots, d\}$,

Proof of Theorem 8.7 (cont)

$$d\langle X^{x,i}, X^{x,k}\rangle_s = \sum_{j=1}^{m} \sigma_{ij}(X_s^x)\sigma_{kj}(X_s^x)ds.$$

We thus see that, if g is the function defined by

$$g(x) = \frac{1}{2}\sum_{i,k=1}^{d} (\sigma\sigma^*)_{ik}(x)\frac{\partial^2 f}{\partial x_i \partial x_k}(x) + \sum_{i=1}^{d} b_i(x)\frac{\partial f}{\partial x_i}(x),$$

the process

$$M_t = f(X_t^x) - f(x) - \int_0^t g(X_s^x)ds$$

is a continuous local martingale. Since f and g are bounded, Proposition 4.7 (ii) shows that M is a martingale. It now follows from Theorem 6.14 that $f \in D(L)$ and $Lf = g$.

Corollary 8.8

Suppose that $(X_t)_{t\geq 0}$ solves $E(\sigma, b)$ on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$. Then $(X_t)_{t\geq 0}$ satisfies the strong Markov property: If T is a stopping time and if Φ is a Borel measurable function from $C(\mathbb{R}_+, \mathbb{R}^d)$ to $\mathbb{R}_+$,

$$\mathbb{E}[1_{\{T<\infty\}}\Phi(X_{T+t}, t \geq 0)|\mathcal{F}_T] = 1_{\{T<\infty\}}\mathbb{E}_{X_T}[\Phi]$$

where, for every $x \in \mathbb{R}^d$, $\mathbb{P}_x$ denotes the law on $C(\mathbb{R}_+, \mathbb{R}^d)$ of an arbitrary solution of $E_x(\sigma, b)$.

Proof

It suffices to apply Theorem 6.17. Alternatively, we could also argue in a similar manner as in the proof of Theorem 8.6, letting the stopping time T play the same role as the deterministic time s in the latter proof, and using the strong Markov property of Brownian motion.

Corollary 8.8

Suppose that $(X_t)_{t\geq 0}$ solves $E(\sigma, b)$ on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$. Then $(X_t)_{t\geq 0}$ satisfies the strong Markov property: If T is a stopping time and if Φ is a Borel measurable function from $C(\mathbb{R}_+, \mathbb{R}^d)$ to $\mathbb{R}_+$,

$$\mathbb{E}[1_{\{T<\infty\}}\Phi(X_{T+t}, t \geq 0)|\mathcal{F}_T] = 1_{\{T<\infty\}}\mathbb{E}_{X_T}[\Phi]$$

where, for every $x \in \mathbb{R}^d$, $\mathbb{P}_x$ denotes the law on $C(\mathbb{R}_+, \mathbb{R}^d)$ of an arbitrary solution of $E_x(\sigma, b)$.

Proof

It suffices to apply Theorem 6.17. Alternatively, we could also argue in a similar manner as in the proof of Theorem 8.6, letting the stopping time T play the same role as the deterministic time s in the latter proof, and using the strong Markov property of Brownian motion.

Markov processes with continuous sample paths that are obtained as solutions of stochastic differential equations are sometimes called diffusion processes (certain authors call a diffusion process any strong Markov process with continuous sample paths in $\mathbb{R}^d$ or on a manifold). Note that, even in the Lipschitz setting considered here, Theorem 8.7 does not completely identify the generator L, but only its action on a subset of the domain $D(L)$: As we already mentioned in Chap. 6, it is often very difficult to give a complete description of the domain. However, in many instances, one can show that a partial knowledge of the generator, such as the one given by Theorem 8.7, suffices to characterize the law of the process. This observation is at the core of the powerful theory of martingale problems, which is developed in the classical book by Stroock and Varadhan.

At least when restricted $C_c^2(\mathbb{R}^d)$, the generator L is a second order differential operator. The stochastic differential equation $E(\sigma, b)$ allows one to give a probabilistic approach (as well as an interpretation) to many analytic results concerning this differential operator, in the spirit of the connections between Brownian motion and the Laplace operator described in the previous chapter. These connections between probability and analysis were an important motivation for the definition and study of stochastic differential equations.

Outline

In this section, we briefly discuss three important examples, all in dimension one. In the first two examples, one can obtain an explicit formula for the solution, which is of course not the case in general.

8.4.1 The Ornstein-Uhlenbeck Process

Let $\lambda > 0$. The (one-dimensional) Ornstein-Uhlenbeck process is the solution of the stochastic differential equation

$$dX_t = dB_t - \lambda X_t dt.$$

This equation is solved by applying Ito's formula to $e^{\lambda t} X_t$,and we get

$$X_t = X_0 e^{-\lambda t} + \int_0^t e^{-\lambda(t-s)} dB_s.$$

Note that the stochastic integral is a Wiener integral (the integrand is deterministic), which thus belongs to the Gaussian space generated by B.

In this section, we briefly discuss three important examples, all in dimension one. In the first two examples, one can obtain an explicit formula for the solution, which is of course not the case in general.

8.4.1 The Ornstein-Uhlenbeck Process

Let $\lambda > 0$. The (one-dimensional) Ornstein-Uhlenbeck process is the solution of the stochastic differential equation

$$dX_t = dB_t - \lambda X_t dt.$$

This equation is solved by applying Ito's formula to $e^{\lambda t} X_t$,and we get

$$X_t = X_0 e^{-\lambda t} + \int_0^t e^{-\lambda(t-s)} dB_s.$$

Note that the stochastic integral is a Wiener integral (the integrand is deterministic), which thus belongs to the Gaussian space generated by B.

First consider the case where $X_0 = x \in \mathbb{R}$. By the previous remark, X is a (non-centered) Gaussian process, whose mean function is $m(t) = \mathbb{E}[X_t] = xe^{-\lambda t}$, and whose covariance function is also easy to compute:

$$K(s,t) = \mathrm{Cov}(X_s, X_t) = \frac{e^{-\lambda|t-s|} - e^{-\lambda(t+s)}}{2\lambda}.$$

It is also interesting to consider the case when X_0 is distributed according to $\mathcal{N}(0, \frac{1}{2\lambda})$. In this case, X is a centered Gaussian process with covariance function

$$\frac{1}{2\lambda} e^{-\lambda|t-s|}.$$

Notice that this is a stationary covariance function. In that case, the Ornstein-Uhlenbeck process X is both a stationary Gaussian process (indexed by $\mathbb{R}_+$) and a Markov process.

8.4.2 Geometric Brownian Motion

Let $\sigma > 0$ and $r \in \mathbb{R}$. The geometric Brownian motion with parameters σ and r is the solution of the stochastic differential equation

$$dX_t = \sigma X_t dB_t + rX_t dt.$$

One solves this equation by applying Ito's formula to $\log X_t$ (say in the case where $X_0 > 0$) and it follows that:

$$X_t = X_0 \exp\left(\sigma B_t + (r - \frac{\sigma^2}{2})t\right).$$

Note in particular that, if the initial value X_0 is (strictly) positive, the solution remains so at every time $t \geq 0$. Geometric Brownian motion is used in the celebrated Black-Scholes model of financial mathematics.

The reason for the use of this process comes from an economic assumption of independence of the successive ratios

$$\frac{X_{t_2} - X_{t_1}}{X_{t_1}}, \frac{X_{t_3} - X_{t_2}}{X_{t_2}}, \ldots, \frac{X_{t_n} - X_{t_{n-1}}}{X_{t_{n-1}}}$$

corresponding to disjoint time intervals: From the explicit formula for X_t, we see that this is nothing but the property of independence of increments of Brownian motion.

8.4.3 Bessel Processes

Let $m \geq 0$. The m-dimensional squared Bessel process is the real-valued process taking non-negative values that solves the stochastic differential equation

$$dX_t = 2\sqrt{X_t}dB_t + mdt. \tag{3}$$

Note that this equation does not fit into the Lipschitz setting studied in this chapter, because the function $\sigma(x) = 2\sqrt{x}$ is not Lipschitz on $\mathbb{R}_+$. However, there exist (especially in dimension one) criteria weaker than our Lipschitz continuity assumptions, which apply to (3) and give the existence and pathwise uniqueness of solutions of (3). For instance, see Exercise 8.14 for a criterion of pathwise uniqueness that applies to (3).

One of the main reasons for studying Bessel processes comes from the following observation. If $d \geq 1$ is an integer and $\beta = (\beta^1, \dots, \beta^d)$ is a d-dimensional Brownian motion, an application of Ito's formula shows that the process

$$|\beta_t|^2 = (\beta_t^1)^2 + \cdots + (\beta_t^d)^2$$

is a d-dimensional squared Bessel process.

Suppose from now on that $m > 0$ and $X_0 = x > 0$. For every $r \geq 0$, define $T_r = \inf\{t \geq 0 : X_t = r\}$. If $r > x$, we have $\mathbb{P}(T_r < \infty) = 1$. To show this, use (3) to see that $X_{t\wedge T_r} = x + m(t \wedge T_r) + Y_{t\wedge T_r}$, where $\mathbb{E}[(Y_{t\wedge T_r})^2] \leq 4rt$. By Markov's inequality, $\mathbb{P}(Y_{t\wedge T_r} > t^{3/4}) \to 0$ as $t \to \infty$, and if we assume that $\mathbb{P}(T_r = \infty) > 0$ the preceding expression for $X_{t\wedge T_r}$ gives a contradiction.

Define, for every $t \in [0, T_0)$,

$$M_t = \begin{cases} (X_t)^{1-\frac{m}{2}}, & m \neq 2, \\ \log(X_t), & , m = 2. \end{cases}$$

It follows from Ito's formula that, for every $\epsilon \in (0, x)$, $M_{t \wedge T_\epsilon}$ is a continuous local martingale. This continuous local martingale is bounded over the time interval $[0, T_\epsilon \wedge T_A]$, for every $A > x$, and an application of the optional stopping theorem (using the fact that $T_A < \infty$ a.s.) gives, if $m \neq 2$,

$$\mathbb{P}(T_\epsilon < T_A) = \frac{A^{1-\frac{m}{2}} - x^{1-\frac{m}{2}}}{A^{1-\frac{m}{2}} - \epsilon^{1-\frac{m}{2}}},$$

and if $m = 2$,

$$\mathbb{P}(T_\epsilon < T_A) = \frac{\log A - \log x}{\log A - \log \epsilon}.$$

Let's finally concentrate on the case $m \geq 2$. Let $\epsilon \downarrow 0$ in the preceding formulas, we obtain that $\mathbb{P}(T_0 < \infty) = 0$. If we let $A \uparrow \infty$, we also get that $\mathbb{P}(T_\epsilon < \infty) = 1$ if $m = 2$ and $\mathbb{P}(T_\epsilon < \infty) = (\epsilon/x)^{(m/2)-1}$ if $m > 2$.

It then follows from the property $\mathbb{P}(T_0 < \infty) = 0$ that the process M_t is well defined for every $t \geq 0$ and is a continuous local martingale. When $m > 2$, M_t takes non-negative values and is thus a supermartingale, which converges a.s. as $t \to \infty$. The limit must be 0, since we already noticed that $\mathbb{P}(T_A < \infty) = 1$ for every $A > x$ and we conclude that X_t converges a.s. to ∞ as $t \to \infty$ when $m > 2$. One can show that the continuous local martingale M_t is not a (true) martingale.