AN INTRODUCTION TO MARKOV PROCESSES AND THEIR APPLICATIONS IN MATHEMATICAL ECONOMICS

UMUT ÇETIN

1. Markov Property

2. Brief review of martingale theory

3. Feller Processes

4. Infinitesimal generators

5. Martingale Problems and Stochastic Differential Equations

Consider the following stochastic differential equation:

$$X_t = X_0 + \int_0^t b(s, X_s)\, ds + \int_0^t \sigma(s, X_s) dW_s \tag{5.1}$$

where W is a r-dimensional Brownian motion, b is a $d \times 1$ drift vector and σ is a $d \times r$ dispersion matrix.

Definition 5.1. *A weak solution of (5.1) is a triple (X, W), $(\Omega, \mathcal{F}, P)$, $(\mathcal{F}_t)$, where*

i) *$(\Omega, \mathcal{F}, P)$ is a probability space, $(\mathcal{F}_t)$ is a filtration of sub-σ-fields of $\mathcal{F}$ satisfying usual conditions;*
ii) *$X = (X_t)$ is adapted to $(\mathcal{F}_t)$ and W is $(\mathcal{F}_t)$-Brownian motion such that (5.1) is satisfied and $P[\int_0^t \left\{|b(s, X_s)| + \sigma_{ij}^2(s, X_s)\right\} ds < \infty] = 1$ for every $t > 0$.*

The probability measure μ on $\mathscr{B}(\mathbb{R}^d)$ defined by $\mu(\Lambda) = P(X_0 \in \Lambda)$ is called the *initial distribution* of the solution.

Definition 5.2. *We say that* uniqueness in the sense of probability law *holds for (5.1) if, for any two weak solutions (X, W), $(\Omega, \mathcal{F}, P)$, $(\mathcal{F}_t)$ and $(\hat{X}, \hat{W})$, $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P})$, $(\hat{\mathcal{F}}_t)$ with the same initial distribution have the same law, i.e. for any $t_1, \dots, t_n$ and $n \geq 1$, $P(X_{t_1} \in E_1, \dots, X_{t_n} \in E_n) = \hat{P}(\hat{X}_{t_1} \in E_1, \dots, \hat{X}_{t_n} \in E_n)$, where $E_i \in \mathscr{B}(\mathbb{R}^d)$.*

Exercise 5.1. *Let X be a weak solution of (5.1) and define*

$$a_{ij}(t, x) = \sum_{k=1}^r \sigma_{ik}(t, x) \sigma_{kj} t, x).$$

Let L_s be the operator defined in Remark 2. Then, for every continuous function $f : \mathbb{R}_+ \times \mathbb{R}^d \mapsto \mathbb{R}$ belonging to $C^{1,2}((0, \infty) \times \mathbb{R}^d)$

$$M_t^f = f(t, X_t) - f(0, X_0) - \int_0^t \left\{ \frac{\partial f}{\partial s} + L_s f(X_s) \right\} ds$$

is a local martingale. If g is another continuous function belonging to $\mathbb{C}^{1,2}((0,\infty)\times\mathbb{R}^d)$, then

$$\langle M^f, M^g\rangle = \sum_{i,j}\int_0^t a_{ij}(s,X_s)\frac{\partial f}{\partial x_i}(s,X_s)\frac{\partial g}{\partial x_j}(s,X_s)\,ds.$$

Furthermore, if $f \in C_K^{1,2}((0,\infty)\times\mathbb{R}^d)$ and σ_{ij} are bounded on the support of f, then M^f is a martingale such that M_t^f is square integrable for each t.

Let $C[0,\infty)^d$ be the space of continuous functions on $\mathbb{R}^d$ and $\mathscr{B}_t = \sigma(X(s); 0 \le s \le t)$ where X is the coordinate process, and set $\mathscr{B} = \vee_{t\ge 0}\mathscr{B}_t$.

Definition 5.3. *A probability measure P on $(C[0,\infty)^d, \mathscr{B})$, under which M^f is a continuous local martingale for every $f \in C^2(\mathbb{R}^d)$ and $P(X_0 \in \Lambda) = \mu(\lambda)$ for any $\Lambda \in \mathscr{B}(\mathbb{R}^d)$ is called a solution to* the local martingale problem for (L_t, μ). *Here the relevant filtration is $(\mathcal{F}_{t+})$ where $(\mathcal{F}_t)$ is the augmentation of $(\mathscr{B}_t)$ with the P-null sets. We will say P^x solves the local martingale problem when μ is the point mass at $x \in \mathbb{R}^d$.*

In view of Exercise 5.1 we see that the existence of a weak solution to (5.1) implies the existence of a solution to the local martingale problem. The next theorem gives the converse.

Theorem 5.1. *Let P be a probability measure on $(C[0,\infty)^d, \mathscr{B})$, under which M^f is a continuous local martingale for the choices $f(x) = x_i$ and $f(x) = x_i x_j$, $1 \le i,j \le d$. Then, there is an r-dimensional Brownian motion $(W_t, \hat{\mathcal{F}}_t)$ defined on an extension $(\hat\Omega, \hat{\mathcal{F}}, \hat P)$ of $(C[0,\infty)^d, \mathscr{B}, P)$ such that (X, W), $(\hat\Omega, \hat{\mathcal{F}}, \hat P)$, $(\hat{\mathcal{F}}_t)$ is a weak solution of (5.1) where X is the coordinate process.*

The proof of the above theorem can be found in Chap. 5 of Karatzas and Shreve. In order to motivate the idea of the proof consider the one dimensional case. Under P, which solves the local martingale problem,

$$M_t := X_t - X_0 - \int_0^t b(s,X_s)\,ds$$

is a continuous local martingale. Then, then exists a Brownian motion, possibly defined in an extension of the underlying probability space, B such that $dM_t = \rho_t dB_t$ for some adapted, measurable ρ. We want to show that $\rho_t = \sigma(t,X_t) := \sqrt{a(t,X_t)}$. To see this first observe that $dX_t^2 = 2X_t dM_t + \{2X_t b(t,X_t) + \rho_t^2\}\,dt$ by applying Ito formula. Moreover, P being solution of the local martingale problem implies

$$X_t^2 - X_0^2 - \int_0^t \{2X_s b(s,X_s) + a(s,X_s)\}\,ds.$$

Comparing two representations yields the claim since any continuous local martingale of finite variation has to be constant.

Equivalence of local martingale problem and weak solutions is summarised in the following:

Corollary 5.1. *The existence of a solution P to the local martingale problem is equivalent to the existence of a weak solution (X, W), $(\hat\Omega, \hat{\mathcal{F}}, \hat P)$, $(\hat{\mathcal{F}}_t)$ to (5.1). The two solutions are related by $P = \hat P X^{-1}$; i.e. X induces the measure P on $(C[0,\infty)^d, \mathscr{B})$.*

Corollary 5.2. *The uniqueness of the solution P to the local martingale problem with initial distribution μ is equivalent to the uniqueness in law for the solutions of (5.1) with the initial distribution μ.*

Because of the difficulty in computing expectations of local martingales, we are interested in the following modification.

Definition 5.4. *A probability measure P on $(C[0,\infty)^d, \mathscr{B})$ is called a solution of the martingale problem associated with (L,μ) if M^f is a martingale for every $C_K^2(\mathbb{R}^d)$ and $P(X_0 \in \Lambda) = \mu(\lambda)$ for any $\Lambda \in \mathscr{B}(\mathbb{R}^d)$.*

Exercise 5.2. *Show that if P is a solution of the martingale problem, then it also solves the local martingale problem. Moreover, if σ_{ij} are locally bounded, then two problems have the same set of solutions.*

Definition 5.5. *We say that the martingale problem of Definition 5.4 is well posed if for every probability measure μ on $\mathscr{B}(\mathbb{R}^d)$ there exists a unique probability measure P^μ, which solves the martingale problem for (L,μ).*

In view of the observed relationship with the martingale problem and weak solutions of (5.1) we have the following result.

Proposition 5.1. *Suppose that σ_{ij} are locally bounded. Then, the martingale problem is well-posed iff for each $x \in \mathbb{R}^d$ there is a unique solution to (5.1) with initial condition $X_0 = x$ in the sense of probability law.*

We now take up the issue of uniqueness for the martingale problem from a different angle.

Definition 5.6. *A collection of D of Borel measurable functions $\phi : \mathbb{R}^d \mapsto \mathbb{R}$ is called a determining class on $\mathbb{R}^d$ if for any two finite measures μ_1 and μ_2 on $\mathscr{B}(\mathbb{R}^d)$, the identity*

$$\int \phi(x)\mu_1(dx) = \int \phi(x)\mu_2(dx)$$

for all $\phi \in D$ implies $\mu_1 = \mu_2$.

Exercise 5.3. *$C_K^\infty(\mathbb{R}^d)$ is a determining class on $\mathbb{R}^d$.*

Lemma 5.1. *Suppose that for every $f \in C_K^\infty$, the Cauchy problem*

$$\begin{aligned} \frac{\partial u}{\partial t} &= Lu; \qquad \text{in } (0,\infty)\times\mathbb{R}^d, \\ u(0,x) &= f(x); \qquad \text{in } \mathbb{R}^d \end{aligned}$$

has a continuous solution u_f in $C^{1,2}((0,\infty)\times\mathbb{R}^d)$ which is bounded on $[0,T]\times\mathbb{R}^d$ for every $T>0$. Then, if P^x and $\tilde{P}^x$ are two solutions of the time homogenous martingale problem associated to L with initial condition $x \in \mathbb{R}^d$, then

$$P^x[X_t \in \Lambda] = \tilde{P}^x[X_t \in \Lambda]$$

for any Borel set Λ.

Proof. For fixed $T > 0$ let $g(t,x) = u_f(T-t,x)$ for $t \le T$. Then, since g is bounded it follows from Ito's formula that $(g(t,X_t)$ is a martingale with respect to P^x and $\tilde{P}^x$. Thus,

$$E^x f(X_T) = E^x g(T,X_T) = g(0,x) = \tilde{E}^x g(T,X_T) = \tilde{E}^x f(X_T).$$

Since f is any function in the determining class $\mathbb{C}_K^\infty$, the claim follows. □

Uniqueness of the marginal distributions remarkably leads to the Markov property of X. Note that one cannot conclude directly that X is always Markov. This needs additional assumptions. Uniqueness of marginal distributions is one such sufficient condition.

Proposition 5.2. *Suppose that for any given $x \in \mathbb{R}^d$ whenever P^x and $\tilde{P}^x$ are solutions to the martingale problem, one has*

$$P^x[X_t \in \Lambda] = \tilde{P}^x[X_t \in \Lambda]$$

for any Borel set Λ. Then X has the Markov property under any P^x.

Proof. Fix a P^x, $r \geq 0$ and let $F \in \mathcal{F}_r$ be such that $P^x(F) > 0$. Define the measures

$$P_1(E) = \frac{E^x[\mathbf{1}_F E^x[\mathbf{1}_E|\mathcal{F}_r]]}{P^x(F)}$$
$$P_1(E) = \frac{E^x[\mathbf{1}_F E^x[\mathbf{1}_E|X_r]]}{P^x(F)}$$

for any $E \in \mathcal{F}$. Set $Y_t(\omega) = X_{t+r}(\omega)$ and note that for any Borel set Λ

$$P_1(Y_0 \in \Lambda) = P_1(Y_0 \in \Lambda) = P^x(X_r \in \Lambda|F).$$

Next if we define

$$\eta(Y) := \left(f(Y_{t_{n+1}}) - f(Y_{t_n}) - \int_{t_n}^{t_{n+1}} Lf(Y_s)ds \right) \Pi_{k=1}^n h_k(Y_{t_k}),$$

for any smooth f with compact support and a collection of bounded measurable functions h_k, then it can be directly checked (and please check) that

$$E_2[\eta(Y)] = E_1[\eta(Y)] = 0.$$

This shows that P_1 and P_2 are solutions to the martingale problem when the coordinate process is given by Y. Due to the uniqueness of marginal distributions we have $E_1[f(Y_t)] = E_2[f(Y_t)]$ for any bounded measurable f. Thus,

$$E^x[\mathbf{1}_F E^x[f(X_{r+t})|\mathcal{F}_r]] = E^x[\mathbf{1}_F E^x[f(X_{r+t})|X_r]],$$

and the arbitrariness of F yields

$$E^x[f(X_{r+t})|\mathcal{F}_r] = E^x[f(X_{r+t})|X_r].$$

□

In view of the above Markov property the uniqueness of marginal distributions under the solutions of the martingale problem leads to the uniqueness of finite dimensional distributions by induction. Thus, we have the following:

Theorem 5.2. *Suppose that the coefficients b and σ are such that for any $f \in C_K^\infty$ the Cauchy problem in Lemma 5.1 has a continuous solution in $C^{1,2}((0,\infty) \times \mathbb{R}^d)$ which is bounded on $[0,T] \times \mathbb{R}^d$ for every $T > 0$. Then, there exists at most one solution to the martingale problem associated to L for any initial condition $x \in \mathbb{R}^d$.*

We see a remarkable duality here: *existence* of a solution to the Cauchy problem implies *uniqueness* for the martingale problem!

Remark 1. *A sufficient condition for the existence of a solution to the Cauchy problem is that the coefficients b_i and a_{ij} are bounded and Hölder continuous, and uniformly elliptic, i.e.*

$$\sum_{i,j} a_{ij}(x)\xi_i\xi_j \geq \lambda \|\xi\|; \, \forall x, \xi \in \mathbb{R}^d \qquad \textit{and some } \lambda > 0.$$

Exercise 5.4. *Suppose that the coefficients b and σ are locally bounded and let P^x be a solution of the time homogeneous martingale problem associated to L with initial condition $x \in \mathbb{R}^d$. Suppose that there exists $f \in \mathbb{C}^2(\mathbb{R}^d)$ such that*

$$Lf(x) + \lambda f(x) \le c,$$

for all $x \in \mathbb{R}^d$ and some $\lambda > 0, c \ge 0$. Then,

$$E^x f(X_t) \le f(x)e^{-\lambda t} + \frac{c}{\lambda}(1 - e^{-\lambda t}); \qquad 0 \le t < \infty, x \in \mathbb{R}^d.$$

6. Connections with partial differential equations

Consider a Feller process X on $\mathbb{R}^d$ with the generator L. We have seen that for any f with compact support that belong to the domain of A one has

$$\frac{d}{dt}P_t f = LP_t f.$$

Now, suppose that A is given by a second order elliptic operator

$$L = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(\cdot)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b_i(\cdot)\frac{\partial}{\partial x_i}.$$

Since $P_t f(x) = E^x[f(X_t)]$, this shows that the function $u(t, x) := E^x[f(X_t)]$ satisfies the partial differential equation

$$u_t = Lu, \qquad u(0, x) = f(x).$$

Moreover, for a continuous c, if we define the transition function

$$Q_t(x, A) = E^x\left[\mathbf{1}_A(X_t)\exp\left(-\int_0^t c(X_S)\,ds\right)\right],$$

for all $A \in \mathscr{B}(\mathbb{R}^d)$, its associated generator is given by $L - c$ with the same domain as L, in view of Exercise 4.2. Therefore, if $v(t, x) := E^x\left[f(X_t)\exp\left(-\int_0^t c(X_S)\,ds\right)\right]$ for some $f \in \mathscr{D}(L)$, we see that

$$v_t = Lv - cv, \qquad v(0, x) = f(x). \tag{6.1}$$

Now, we are interested in the inverse problem. Namely, if we are given a partial differential equation, can we represent its solution as an expectation? In order to obtain such a representation, clearly we need a form of uniqueness result for the solutions of the associated PDEs. In PDE Theory there are various uniqueness results for a class of initial conditions and coefficients. These results ensures the uniqueness of the solution of the pde *only* in a class of continuous functions satisfying certain growth conditions. Thus, in order to obtain a representation result, say, for the equation in (6.1), it is not enough just to choose nice coefficients and initial condition, but one must ensure that the function defined by $E^x\left[f(X_t)\exp\left(-\int_0^t c(X_S)\,ds\right)\right]$ belongs to the family of functions in which the uniqueness holds.

To this end we assume that there exists a unique weak solution to

$$X_t^{(s,x)} = x + \int_s^t b(r, X_r^{(s,x)})\,dr + \int_0^t \sigma(r, X_r^{(s,x)})dW_r \tag{6.2}$$

for any $(s,x) \in \mathbb{R}_+ \times \mathbb{R}^d$, and that b_i and σ_{ij} are continuous and satisfy the linear growth condition

$$\|b(t,x)\|^2 + \|\sigma(t,x)\|^2 \le K(1+\|x\|^2) \tag{6.3}$$

for any $(t,x) \in \mathbb{R}_+ \times \mathbb{R}^d$.

Exercise 6.1. *Suppose that b and σ do not depend on t and*

$$\sum_{i,j} a_{ij}(x)\xi_i\xi_j \ge \lambda\|\xi\|;\ \forall x \in D, \xi \in \mathbb{R}^d \qquad \textit{and some } \lambda > 0$$

for some opend and bounded *domain D. Let u be a solution of*

$$Lu - ku = -g;\ \textit{in } D$$

with the boundary condition $u = f$ on the boundary of D, where k is a positive continuous function and, f and g are continuous functions. Further suppose that $E^x\tau_D < \infty$ where $\tau_d = \inf\{t \ge 0 : X_t \notin D\}$. Under the assumptions set out above show that for every $x \in D$

$$u(x) = E^x\left[f(X_{\tau_D})\exp\left(-\int_0^{\tau_D} k(X_s)ds\right) + \int_0^{\tau_D} g(X_s)\exp\left(-\int_0^t k(X_s)ds\right)dt\right].$$

(Hint: Consider the process

$$N_t = u(X_{t\wedge\tau_D})\exp\left(-\int_0^{t\wedge\tau_D} k(X_s)ds\right) + \int_0^{t\wedge\tau_D} g(X_s)\exp\left(-\int_0^s k(X_r)dr\right)ds,$$

and first show that it is uniformly integrable martingale under P^x.)

The next result is an analogue of Feynman-Kac formula in this setting. Its proof can be found in Chapter 5 of Karatzas and Shreve.

Suppose that $f : \mathbb{R}^d \mapsto \mathbb{R}, g : [0,T] \times \mathbb{R}^d \mapsto \mathbb{R}$ and k" $[0,T] \times \mathbb{R}^d \mapsto [0,\infty)$ are continuous and satisfy for some $K > 0, \lambda \ge 1$,

$$\begin{aligned} |f(x)| &\le K(1+\|x\|^{2\lambda}) \qquad \text{or} \qquad f \ge 0; \\ |g(t,x)| &\le K(1+\|x\|^{2\lambda}) \qquad \text{or} \qquad g \ge 0. \end{aligned}$$

Then, under the assumptions of this section we have

Theorem 6.1. *Suppse that v is continuous and of class $\mathbb{C}^{1,2}((0,T) \times \mathbb{R}^d)$ and satisfies the Cauchy problem*

$$\begin{aligned} -\frac{\partial v}{\partial t} + kv - g &= L_t v,\ \textit{in } [0,T) \times \mathbb{R}^d, \\ v(T,x) &= f(x),\ x \in \mathbb{R}^d. \end{aligned}$$

Suppose further that

$$\sup_{0\le t\le T} |v(t,x)| \le M(1+\|x\|^{2\mu}),\ x \in \mathbb{R}^d,$$

for some $M > 0$ and $\mu \ge 1$. Then,

$$v(t,x) = E^x\left[f(X_T)\exp\left(-\int_t^T k(u,X_u)du\right) + \int_t^T g(s,X_s)\exp\left(-\int_t^s k(r,X_r)dr\right)ds\right].$$

Definition 6.1. *A fundamental solution of the second-order partial differential equation*

$$-\frac{\partial u}{\partial t} + ku = L_t u \tag{6.4}$$

is a nonnegative function $G(t, x; u, y)$ *such that for every* $f \in C_K(\mathbb{R}^d)$ *the function*

$$u(t, x) = \int_{\mathbb{R}^d} G(t, x; u, y) f(y) dy$$

is bounded, of class $\mathbb{C}^{1,2}$*, satisfies the PDE above, and*

$$\lim_{t \uparrow u} u(t, x) = f(x) \; x \in \mathbb{R}^d.$$

Suppose that there exists a fundamental solution for (6.4), and for the moment suppose that $k \equiv 0$. Then, the pde (6.4) with the boundary condition $u(T, x) = f$ is given by

$$u(t, x) = \int_{\mathbb{R}^d} G(t, x; T, y) f(y) dy.$$

If we further suppose that the coefficients of the PDE are well-behaved so that the stochastic representation holds for the solutions of (6.4), we see that

$$u(t, x) = E^x[f(X_T)].$$

However, since $\mathbb{C}_K^\infty$ is a determining class, this lets us conclude that $G(t, x; u, y)$ is the transition density for the Markov process X. By the same reasoning when k is not necessarily 0, G can be viewed as the transition density of X killed at rate k. For further details of fundamental solutions of partial differential equations considered here and their solutions in general see *Partial differential equations of parabolic type* by A. Friedman.

Department of Statistics, London School of Economics and Political Science, 10 Houghton st, London, WC2A 2AE, UK

E-mail address: u.cetin@lse.ac.uk