9831 Probability and Stochastic Processes for Finance, Fall 2016

Lecture 10

Anja Richter

Department of Mathematics

Outline

- Stochastic differential equations
- Connections to partial differential equations

1 Stochastic differential equations

Let $(\Omega, \mathcal{F}_t, \mathbb{P})$ be a filtered probability space and W_t a Brownian motion defined on it.

A **stochastic differential equation (SDE)** driven by the Brownian motion W_t, is an equation of the form

$$dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dW_t, \quad X_0 = x \qquad (*)$$

or in integral form

$$X_t = x + \int_0^t \mu(X_s, s)ds + \int_0^t \sigma(X_s, s)dW_s,$$

where $\mu(x, t)$ and $\sigma(x, t)$ are deterministic functions of x and t.

As usual, $\mu(x, t)$ is referred to as the **drift part** and $\sigma(x, t)$ the **diffusion part**. We refer to $X_0 = 0$ as **initial condition**.

1.1 Existence and uniqueness of a solution

Definition

For a given initial condition $X_0 = x$, a **solution** to the SDE

$$dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dW_t$$

is a stochastic process X_t with continuous sample paths and satisfying:

- $X_0 = x$ almost surely
- X_t is adapted to the filtration generated by the Brownian motion W_t
- the integral version of the equation holds almost surely

$$X_t = x + \int_0^t \mu(X_s, s)ds + \int_0^t \sigma(X_s, s)dW_s, \quad \text{for } 0 \leq t \leq T.$$

Remark:

- Of course we need that both, the stochastic and the usual integral exist.
- Often in applications, we need to start at time t from a given point x and find X_u, $u \geq t$, such that

$$X_u = x + \int_t^u \mu(X_s, s)ds + \int_t^u \sigma(X_s, s)dW_s, \quad \text{for } t \leq u.$$

Theorem

The SDE

$$dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dW_t, \quad X_0 = x.$$

has a unique solution, if the coefficients $\mu(x, t)$ and $\sigma(x, t)$ satisfy

- uniformly Lipschitz: there is a Lipschitz constant K such that

$$|\mu(x, t) - \mu(y, t)| + |\sigma(x, t) - \sigma(y, t)| \leq K|x - y|.$$

- linear growth: there is a constant such that

$$|\mu(x, t)| + |\sigma(x, t)| \leq G(1 + |x|).$$

Remarks

- The proof of the theorem is based on the technique of **Picard iteration** and **contraction mapping principle**.
- Lipschitz $\leftrightsquigarrow$ uniqueness
- Linear growth $\leftrightsquigarrow$ nonexplosive

1.2 Examples

Explosive diffusion

Consider the SDE
$$dX_t = X_t^3 dt - X_t^2 dW_t, \quad X_0 = 1.$$
We try to find a solution X_t in the form $f(t, W_t)$. Then
$$\begin{aligned} dX_t &= df(t, W_t) \\ &= (f_t + \frac{1}{2} f_{xx}\big) dt + f_x dW_t, \end{aligned}$$
and hence f has to satisfy
$$f_x = -f^2 \quad \text{and} \quad f_t + \frac{1}{2} f_{xx} = f^3.$$
From $f_x = -f^2$ we get
$$f(t, x) = \frac{1}{x + C(t)}.$$
Substituting this into the second equation gives
$$\begin{gathered} f_t + \frac{1}{2} f_{xx} = f^3 \\ \frac{C'(t)}{(x + C(t))^2} + \frac{1}{2} \frac{2}{(x + C(t))^3} = \frac{1}{(x + C(t))^3}. \end{gathered}$$
Hence, $C'(t) = 0$, i.e. $C(t) \equiv$ constant and $f(t, x) = \dfrac{1}{x + C}$.

Finally, we obtain
$$\begin{aligned} X_0 &= f(0, 0) = \frac{1}{C} = 1 \Longrightarrow C = 1 \\ X_t &= f(t, W_t) = \frac{1}{1 + W_t}. \end{aligned}$$
Note that with probability 1, the Brownian motion W_t hits level -1, i.e. $P(\tau_{-1} < \infty) = 1$.

Thus, the solution X_t exists only up to the 'explosion time' τ_{-1}.

Nonunique solution

The SDE

$$dX_t = 3\sqrt[3]{X_t^2}\,dW_t + 3\sqrt[3]{X_t}\,dt, \quad X_0 = 0.$$

has infinitely many solutions which are given by $\varphi(W_t; a)$, where $\varphi(x; a) = (x-a)^3 1_{[a,\infty)}(x)$ for any $a > 0$.

Note that

- $\varphi'(x; a) = 3\varphi(x; a)^{\frac{2}{3}}$
- $\varphi''(x; a) = 6\varphi(x; a)^{\frac{1}{3}}$

1.3 Markov property

Note that in the definition of SDEs, the only randomness that appears is that of the Brownian motion W_t and the solution X_t itself. In particular, we have that μ and σ are deterministic.

This ensures that the solutions to SDEs have the Markov property, which we will justify heuristically in the following.

We can simulate the solution X_t, $0 \leq t \leq T$ to an SDE:

Suppose $X_0 = x$ is given.

Consider a small increment of time Δ, then according to $(*)$ we get

$$\begin{aligned} X_\Delta &\approx x + \mu(x,0)\Delta + \sigma(x,0)W_\Delta \\ &= x + \mu(x,0)\Delta + \sigma(x,0)\sqrt{\Delta}Z_1, \qquad (Z_1 \sim N(0,1)). \end{aligned}$$

Moreover, we have

$$\begin{aligned} X_{(k+1)\Delta} &\approx X_{k\Delta} + \mu(X_{k\Delta},\Delta)\Delta + \sigma(X_{k\Delta},\Delta)(W_{(k+1)\Delta} - W_{k\Delta}) \\ &= X_{k\Delta} + \mu(X_{k\Delta},\Delta)\Delta + \sigma(X_{k\Delta},\Delta)\sqrt{\Delta}Z_{k+1}, \end{aligned}$$

where $Z_1, \ldots, Z_n$ are i.i.d. $N(0,1)$ random variables.

If we can justify passing to the limit (in some sense) as $\Delta \to 0$, then the limiting process gives us a solution to $(*)$.

Moreover, our procedure suggests that the limiting process would be memoryless (i.e. a Markov process), as given $X_{k\Delta}$ we can simulate $X_{(k+1)\Delta}$ without knowing $X_0, X_\Delta, \ldots, X_{(k-1)\Delta}$.

More rigorously, the following theorem holds:

Theorem

Let $t \in [0, T]$, h be a Borel measurable function and set

$$u(x,t) = \mathbb{E}_{t,x}[h(X_T)],$$

then for $t \in [0, T]$,

$$\mathbb{E}[h(X_T) \mid \mathcal{F}_t] = u(X_t, t),$$

i.e. solutions to SDEs are Markov processes.

Remark:

Note that calculating $u(x,t) = \mathbb{E}_{t,x}[h(X_T)]$ means

- take $X_t = x$
- solve $(*)$ for initial condition $X_t = x$ and $u \in [t, T]$
- compute the expectation of $h(X_T)$
- repeat for all $t \in [0, T]$, $x \in \mathbb{R}$
- get a deterministic function $u(x,t)$

2 Connection to partial differential equations

SDEs provide a way to numerically solve second order parabolic partial differential equations (PDEs) by Monte Carlo simulation.

The key point is a stochastic representation of the solution to PDEs which we develop in the following.

Let X_t be the diffusion process driven by

$$dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dW_t,$$

where μ and σ are deterministic variables of space x and time t.

In the following, we shall suppress the dependence on x, t of μ and σ for notational simplicity.

Theorem

Let $u = u(x, t)$ be the solution to the terminal value problem

$$\begin{aligned} & u_t + \frac{\sigma^2}{2} u_{xx} + \mu u_x = 0, \quad t < T, \\ & u(x, T) = h(x). \end{aligned}$$

Then u has the representation

$$u(x, t) = \mathbb{E}_{t,x}\left[h(X_T)\right],$$

where $\mathbb{E}_{t,x}[\cdot]$ denotes the conditional expectation $\mathbb{E}[\cdot | X_t = x]$.

Proof:

Applying Ito's formula to $u(X_t, t)$ yields

$$\begin{aligned} u(X_T, T) - u(X_t, t) &= \int_t^T \sigma(X_s, s) u_x(X_s, s) dW_s + \int_t^T \left[u_t + \frac{\sigma^2}{2} u_{xx} + \mu u_x\right] ds \\ &= \int_t^T \sigma(X_s, s) u_x(X_s, s) dW_s, \end{aligned}$$

since u satisfies the PDE

$$u_t + \frac{\sigma^2}{2} u_{xx} + \mu u_x = 0.$$

Thus taking conditional expectation $\mathbb{E}_{t,x}[\cdot]$ on both sides and taking into account the terminal condition $u(x, T) = h(x)$ we end up with

$$u(x, t) = \mathbb{E}_{t,x}[h(X_T)].$$

Example

Find the solution to the PDE

$$\begin{aligned} u_t(x,t) + \frac{\sigma^2}{2} u_{xx}(x,t) &= 0 \\ u(x,T) &= x^2, \end{aligned}$$

where σ is a constant.

From the above theorem we immediately have

$$u(x,t) = \mathbb{E}_{t,x}[X_T^2],$$

where

$$\begin{aligned} dX_s &= 0 \cdot ds + \sigma dW_s \\ X_t &= x. \end{aligned}$$

Solving this equation gives

$$X_T = x + \sigma(W_T - W_t),$$

so $X_T \sim N(x, \sigma^2(T-t))$.

Thus we have

$$\begin{aligned} u(x,t) &= \mathbb{E}_{t,x}[X_T^2] \\ &= \mathrm{Var}[X_T] + (\mathbb{E}X_T)^2 \\ &= \sigma^2(T-t) + x^2. \end{aligned}$$

2.1 The Feynman-Kac formula

Let $u = u(x,t)$ be the solution to the terminal value problem

$$\begin{aligned} u_t + \frac{\sigma^2}{2} u_{xx} + \mu u_x &= V(x,t)u, \quad t < T, \\ u(x,T) &= h(x). \end{aligned}$$

Then u has the representation

$$u(x,t) = \mathbb{E}_{t,x}\left[e^{-\int_t^T V(X_s,s)ds} h(X_T)\right],$$

where $\mathbb{E}_{t,x}[\cdot]$ denotes the conditional expectation $\mathbb{E}[\cdot | X_t = x]$. This is the celebrated **Feynman-Kac formula**.

Note:

We assume that some mild integrability conditions are satisfied so that all integrals make sense.

Proof:

Applying Ito's formula to $u(X_t,t)e^{-\int_0^t V(X_s,s)ds}$ yields

$$\begin{aligned}
&d\left[u(X_t,t)e^{-\int_0^t V(X_s,s)ds}\right] \\
&= e^{-\int_0^t V(X_s,s)ds}\left[u_t(X_t,t)dt - V(X_t,t)u(X_t,t)dt + u_x(X_t,t)dX_t + \frac{1}{2}u_{xx}(X_t,t)d[X]_t\right] \\
&= e^{-\int_0^t V(X_s,s)ds}\left[\left\{u_t(X_t,t) + \frac{\sigma^2(X_t,t)}{2}u_{xx}(X_t,t) + \mu u_x(X_t,t) - V(X_t,t)u(X_t,t)\right\}dt\right. \\
&\qquad\qquad\qquad \left. + \sigma(X_t,t)u_x(X_t,t)dB_t\right] \\
&= e^{-\int_0^t V(X_s,s)ds}\sigma(X_t,t)u_x(X_t,t)dB_t
\end{aligned}$$

since u satisfies the PDE

$$u_t + \frac{\sigma^2}{2}u_{xx} + \mu u_x = V(x,t)u.$$

In integral form

$$u(X_T,T)e^{-\int_0^T V(X_s,s)ds} - u(X_t,t)e^{-\int_0^t V(X_s,s)ds} = \int_t^T e^{-\int_0^\tau V(X_s,s)ds}\sigma(X_\tau,\tau)u_x(X_\tau,\tau)dB_\tau$$

therefore, by dividing on both sides the term $e^{-\int_0^t V(X_s,s)ds}$ we have

$$u(X_T,T)e^{-\int_t^T V(X_s,s)ds} - u(X_t,t) = \int_t^T e^{-\int_t^\tau V(X_s,s)ds}\sigma(X_\tau,\tau)u_x(X_\tau,\tau)dB_\tau.$$

Thus, taking conditional expectation $\mathbb{E}_{t,x}[\cdot]$ on both sides and taking into account the terminal condition $u(x,T) = h(x)$ we end up with

$$u(x,t) = \mathbb{E}_{t,x}\left[e^{-\int_t^T V(X_s,s)ds}h(X_T)\right].$$

2.2 Multidimensional Feynman-Kac formula

Let X_t be the solution to the multidimensional SDE

$$dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dW_t, \quad X_0 = x, \qquad (**)$$

where

- W_t is an n-dimensional Brownian motion,
- $\mu(x,t)$ is d-dimensional deterministic function on $\mathbb{R}^d \times [0,\infty)$ and
- $\sigma(x,t)$ is a deterministic $d \times n$ matrix for all $(x,t) \in \mathbb{R}^d \times [0,\infty)$.

Let $u = u(x,t)$ be the solution to the terminal value problem

$$u_t + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}(\sigma\sigma^T)_{ij}(x,t)u\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{j=1}^{d}\mu_j(x,t)\frac{\partial}{\partial x_j}u = V(x,t)u, \quad t < T,$$
$$u(x,T) = h(x).$$

Then u has the representation

$$u(x,t) = \mathbb{E}_{t,x}\left[e^{-\int_t^T V(X_s,s)ds}h(X_T)\right].$$

This is the multidimensional Feynman-Kac formula.

Example

Suppose that the stock price under the risk neutral measure $\widetilde{\mathbb{P}}$ satisfies

$$\begin{aligned} dS_t &= rS_t dt + \sqrt{V_t} S_t d\tilde{B}_1(t) \\ dV_t &= (a - bV_t)dt + \sigma\sqrt{V_t}\, d\tilde{B}_2(t), \end{aligned}$$

where $a, b, \sigma > 0$ and $d[B_1, B_2]_t = \rho dt$ with $\rho \in (-1, 1)$.

According to the risk-neutral pricing formula, the price of a derivative with payoff $h(S_T, V_T)$ is

$$P_t = \widetilde{\mathbb{E}}[e^{-r(T-t)} h(S_T, V_T) \mid \mathcal{F}_t].$$

Because solutions to SDEs are Markov processes, there exists a function $u(x, v, t)$ such that $P_t = u(S_t, V_t, t)$.

Find the PDE satisfied by $u(x, v, t)$.

We use the above theorem to identify:

- $V(x, v, t) \equiv r$
- $\mu(x, v, t) = \begin{pmatrix} rx \\ a - bv \end{pmatrix}$
- Finding $\sigma(x, v, t)$ to compute $\sigma\sigma^T$ involves decorrelating the two Brownian motions

  $$\begin{aligned} d\tilde{B}_1(t) &= d\tilde{W}_1(t) \\ d\tilde{B}_2(t) &= \rho d\tilde{W}_1(t) + \sqrt{1-\rho^2}\, d\tilde{W}_2(t), \end{aligned}$$

  where $(\tilde{W}_1(t), \tilde{W}_2(t))$ is a standard 2-dimensional Brownian motion under $\widetilde{\mathbb{P}}$.

  Hence we have

  $$\begin{aligned} dS_t &= rS_t dt + \sqrt{V_t} S_t d\tilde{W}_1(t) \\ dV_t &= (a - bV_t)dt + \sigma\sqrt{V_t}\left(\rho d\tilde{W}_1(t) + \sqrt{1-\rho^2}\, d\tilde{W}_2(t)\right), \end{aligned}$$

  and thus

  $$\sigma(x, v, t) = \begin{pmatrix} \sqrt{v}x & 0 \\ \rho\sigma\sqrt{v} & \sqrt{1-\rho^2}\sigma\sqrt{v} \end{pmatrix},$$

  which gives

  $$\sigma\sigma^T(x, v, t) = \begin{pmatrix} vx^2 & \sigma x v\rho \\ \sigma x v \rho & \sigma^2 v \end{pmatrix},$$

The multidimensional Feynman-Kac formula then gives

$$\partial_t u + rx\partial_x u + (a - bv)\partial_v u + \frac{1}{2} v x^2 \partial_{xx} u + \sigma\rho v x \partial_{xv} u + \frac{1}{2}\sigma^2 v \partial_{vv} u = ru,$$

with terminal condition $u(x, v, t) = h(x, v)$.

2.3 Adding a nonhomogeneous term

Let $u = u(x, t)$ be the solution to the terminal value problem

$$u_t + \frac{\sigma^2}{2} u_{xx} + \mu u_x = f(x, t), \quad t < T,$$
$$u(x, T) = h(x).$$

Then u has the representation

$$u(x, t) = \mathbb{E}_{t,x}\left[h(X_T) - \int_t^T f(X_\tau, \tau) d\tau\right],$$

where $\mathbb{E}_{t,x}[\cdot]$ denotes the conditional expectation $\mathbb{E}[\cdot | X_t = x]$.

Proof:

Applying Ito's formula to $u(X_t, t) - \int_0^t f(X_\tau, \tau) d\tau$ yields

$$d\left[u(X_t, t) - \int_0^t f(X_\tau, \tau) d\tau\right]$$
$$= \left[u_t(X_t, t) + \frac{\sigma^2(X_t, t)}{2} u_{xx}(X_, t) + \mu(X_t, t) u_x(X_t, t) - f(X_t, t)\right] dt + \sigma(X_t, t) u_x(X_t, t) dB_t$$
$$= \sigma(X_t, t) u_x(X_t, t) dB_t$$

since u satisfies the PDE

$$u_t + \frac{\sigma^2}{2} u_{xx} + \mu u_x = f(x, t).$$

In integral form

$$u(X_T, T) - u(X_t, t) - \int_t^T f(X_\tau, \tau) d\tau = \int_t^T \sigma(X_\tau, \tau) u_x(X_\tau, \tau) dB_\tau.$$

Thus taking conditional expectation $\mathbb{E}_{t,x}[\cdot]$ on both sides and taking into account the terminal condition we end up with

$$u(x, t) = \mathbb{E}_{t,x}\left[h(X_T) - \int_t^T f(X_\tau, \tau) d\tau\right].$$

2.4 Backward second order parabolic PDEs

Finally, we have the stochastic representation for *any* backward second order parabolic linear PDE with terminal condition as follows.

Theorem

Let $u = u(x,t)$ be the solution to the terminal value problem

$$\begin{aligned} &u_t + \frac{\sigma^2}{2}u_{xx} + \mu u_x = V(x,t)u + f(x,t), \quad t < T, \\ &u(x,T) = h(x). \end{aligned}$$

Then u has the representation

$$u(x,t) = \mathbb{E}_{t,x}\left[e^{-\int_t^T V(X_s,s)ds}h(X_T) - \int_t^T e^{-\int_t^\tau V(X_s,s)ds} f(X_\tau,\tau)d\tau\right],$$

where $\mathbb{E}_{t,x}[\cdot] = \mathbb{E}[\cdot|X_t = x]$ is the conditional expectation.

2.5 Backward Kolmogorov equation

Let $p(s,y|t,x)$ be the transition density from x at time t to y at time s ($t < s$) of the diffusion process

$$dX_t = \mu(X_t,t)dt + \sigma(X_t,t)dB_t,$$

i.e.

$$\mathbb{P}(X_s \in B \,|\, X_t = x) = \int_B p(s,y|t,x)dy$$

In other words, given that $X_t = x$, for a given $s > t$, X_s is a random variable, whose distribution depends on μ and σ as well as (t,x).

The random variable X_s has the density $p(s,y|t,x)$, where s,t,x are fixed parameters, i.e. the density is a function of y.

Note that

- the variables (t,x) are usually referred to as *backward variables* as they correspond to the starting point $X_t = x$.
- the variables (s,y) are usually referred to as *forward variables* as y is the location of the process at a future time s.

Question: How to find $p(s,y|t,x)$?

Example

Consider the heat equation

$$u_t + \frac{1}{2}u_{xx} = 0$$
$$u(x,T) = h(x).$$

We have the following solution u: $u(x,t) = \mathbb{E}_{t,x}[h(X_T)]$, where X_t is given as

$$dX_t = dW_t, \quad X_0 = x,$$

as $\mu \equiv 0$ and $\sigma^2 \equiv 1$.

Thus $X_u = x + W_u - W_t$ for $u \in [t,T]$, i.e.

$$u(x,t) = \mathbb{E}_{t,x}[h(x + W_T - W_t)] = \int_{-\infty}^{\infty} h(x + \sqrt{T-t}z)\frac{1}{\sqrt{2\pi}}e^{-z^2/2}dz. \qquad (1)$$

On the other hand, from the perspective of $p(T,y|t,x)$, we have

$$u(x,t) = \mathbb{E}_{t,x}[h(X_T)] = \int_{-\infty}^{\infty} h(y)p(T,y|t,x)dy. \qquad (2)$$

Setting $y = x + \sqrt{T-t}z$ in (1), we get that

$$p(T,y|t,x) = \frac{1}{\sqrt{2\pi(T-t)}}e^{-\frac{(y-x)^2}{2(T-t)}}.$$

This is the transition probabilility density of a standard Brownian motion.

Note that $p(T,y|t,x)dy$ is the probability that the Brownian motion started at x at time t will be in $(y, y+dy)$ at time T.

Recall that if u is the solution to the PDE

$$u_t + \frac{\sigma^2}{2} u_{xx} + \mu u_x = 0$$

with terminal condition $u(x,T) = h(x)$, then u has the representation

$$u(x,t) = \mathbb{E}_{t,x}[h(X_T)].$$

Now set $h(X_T) = 1_B(X_T)$, then the representation can be written as

$$u(x,t) = \mathbb{E}_{t,x}[1_B(X_T)] = \int 1_B(y) p(T,y|t,x)dy = \int_B p(T,y|t,x)dy.$$

Therefore, for fixed s and y, the transition probability $\mathbb{P}(X_T \in B \mid X_t = x)$ as a function of t, x satisfies the same PDE as u.

Using basically the same reasoning as above, the transition density $p(s,y|t,x)$, for fixed s and y and as a function of t, x satisfies the same PDE as u, i.e.,

$$\frac{\partial p}{\partial t} + \frac{1}{2}\sigma^2(x,t)\frac{\partial^2 p}{\partial x^2} + \mu(x,t)\frac{\partial p}{\partial x} = 0, \qquad t < s,$$

with terminal condition

$$p(s,y|s,x) = \delta_y(x),$$

which is also referred to as the **backward Kolmogorov equation**.

2.6 Forward Kolmogorov equation

On the other hand, for fixed t and x, the transition density p as a function of s, y satisfies the **Kolmogorov forward equation** (also called the **Fokker-Planck equation**).

We derive it as follows. By the Markov property of the process X_u, we have

$$p(\tau, \xi | t, x) = \int p(s, y | t, x) p(\tau, \xi | s, y) dy, \qquad \text{for } t < s < \tau.$$

Differentiate the equation on both sides with respect to s

$$\begin{aligned}
0 &= \int \frac{\partial p}{\partial s}(s, y|t, x) p(\tau, \xi|s, y) dy + \int p(s, y|t, x) \frac{\partial p}{\partial t}(\tau, \xi|s, y) dy \\
&= \int \frac{\partial p}{\partial s}(s, y|t, x) p(\tau, \xi|s, y) dy - \int p(s, y|t, x) \left[\frac{1}{2}\sigma^2(y, s) p_{xx} + \mu(y, s) p_x\right] dy \\
&= \int \frac{\partial p}{\partial s}(s, y|t, x) p(\tau, \xi|s, y) dy - \int \left[\frac{1}{2}\frac{\partial^2}{\partial y^2}(\sigma^2(y, s)p) - \frac{\partial}{\partial y}(\mu(y, s)p)\right] p(\tau, \xi|s, y) dy \\
&= \int \left[p_s - \frac{1}{2}(\sigma^2 p)_{yy} + (\mu p)_y\right] p(\tau, \xi|s, y) dy.
\end{aligned}$$

The above equation holds for all τ, ξ, therefore p (for fixed t, x) as a function of s, y satisfies the forward equation

$$\frac{\partial p}{\partial s} = \frac{1}{2}\frac{\partial^2}{\partial y^2}\left[\sigma^2(y, s)p\right] - \frac{\partial}{\partial y}\left[\mu(y, s)p\right], \qquad s > t,$$

with initial condition

$$p(t, y|t, x) = \delta_x(y).$$

2.7 Backward stochastic differential equation (BSDE)

Notice that the stochastic representations we obtained so far are all for second order *linear* parabolic PDEs.

What if the PDE is nonlinear? Does there exist an analogue for nonlinear PDEs?

Backward stochastic differential equations (BSDE) provide a stochastic representation of a certain type of second order *nonlinear* parabolic equations such as

$$u_t + \frac{1}{2}u_{xx} + g(t, u, u_x) = 0,$$

where g is a certain function which introduces nonlinearity to the equation.

Definition

A BSDE is an differential equation of the form

$$-dY_t = g(t, Y_t, Z_t)dt - Z_t dW_t$$

accompanied with a terminal condition

$$Y_T = \xi.$$

Here ξ is an $\mathcal{F}_T$-measurable random variable. Equivalently, in integral form, a BSDE reads

$$\xi_T - Y_t = -\int_t^T g(s, Y_s, Z_s)ds + \int_t^T Z_s dW_s.$$

Note

- g is usually referred to as the *generator* of the BSDE.
- The solution to the BSDE is the pair of process (Y_t, Z_t).
- The Z_t process is required so that Y_t is adapted.

2.8 Link to nonlinear PDEs

Consider the nonlinear PDE

$$u_t + \frac{1}{2}u_{xx} = g(t, x, u, u_x)$$

with terminal condition

$$u(x, T) = \phi(x).$$

To see why BSDEs provide stochastic representations for nonlinear PDEs, let $u(t, x)$ be a solution to the PDE and define $Y_t = u(t, W_t)$ and $Z_t = u_x(t, W_t)$, where W_t is the driving Brownian motion.

Applying Ito's formula to $Y_t = u(t, W_t)$ we have

$$\begin{aligned} dY_t &= \left(u_t + \frac{1}{2}u_{xx}\right) dt + u_x(t, W_t) dW_t \\ &= -g(t, W_t, u(t, W_t), u_x(t, W_t))dt + u_x(t, W_t)dW_t. \end{aligned}$$

Hence,

$$\begin{aligned} u(T, W_T) - u(t, W_t) &= -\int_t^T g(s, W_s, u, u_x)ds + \int_t^T u_x dW_s \\ \implies \phi(W_T) - Y_t &= -\int_t^T g(s, W_s, Y_s, Z_s)ds + \int_t^T Z_s dW_s. \end{aligned}$$

In other words, the solution (Y_t, Z_t) to BSDE is given by the solution to its associated nonlinear PDE, i.e., $(Y_t, Z_t) = (u(t, W_t), u_x(t, W_t))$. The terminal condition for the BSDE is $Y_T = \phi(W_T)$.

On the other hand, we obtain the stochastic representation for the nonlinear PDE as

$$u(t, x) = \mathbb{E}\left[\phi(W_T) + \int_t^T g(s, W_s, Y_s, Z_s)dt \,\middle|\, W_t = x\right].$$

We are cheating a bit though: In order to evalute the exepectation on the right hand side, we will need to know how to simulate the processes Y_t and Z_t, which are given by u and u_x!

2.9 Forward backward stochastic differential equation (FBSDE)

By the same token, FBSDE provide a stochastic representation for nonlinear PDE associated with general processes.

Definition

An FBSDE is an equation of the form

$$\begin{aligned} dX_t &= \mu(X_t, t)dt + \sigma(X_t, t)dW_t, \\ -dY_t &= g(t, X_t, Y_t, Z_t)dt - Z_t dW_t \end{aligned}$$

with terminal condition $Y_T = \varphi(X_T)$.

Here $\sigma = \sigma(t, x)$ and $\mu = \mu(t, x)$ are smooth, deterministic functions. The nonlinear PDE associated with the FBSDE is

$$u_t + \frac{\sigma^2}{2} u_{xx} + \mu u_x + g(t, x, u, \sigma u_x) = 0$$

with terminal condition $u(T, x) = \varphi(x)$.

2.10 Stochastic representation with Itô process

Let $Y_t = u(t, X_t)$ and $Z_t = \sigma(t, X_t)u_x(t, X_t)$. Apply Ito's formula to $Y_t = u(t, X_t)$

$$\begin{aligned} dY_t &= du(t, X_t) \\ &= \left[u_t + \frac{\sigma^2}{2}u_{xx}(t, X_t) + \mu u_x(t, X_t)\right] dt + \sigma(t, X_t)u_x(t, X_t)dW_t \\ &= -g(t, X_t, u(t, X_t), \sigma u_x(t, X_t))dt + \sigma(t, X_t)u_x(t, X_t)dW_t \\ &= -g(t, X_t, Y_t, Z_t)dt + Z_t dW_t \end{aligned}$$

Or equivalently, in integral form

$$\varphi(X_T) - Y_t = -\int_t^T g(s, X_s, Y_s, Z_s)ds + \int_t^T Z_s dW_s.$$

Moreover, the conditional expectation

$$u(t, x) = \mathbb{E}\left[\varphi(X_T) + \int_t^T g(s, X_s, Y_s, Z_s)ds \middle| X_t = x\right]$$

is a solution to the PDE

$$u_t + \frac{\sigma^2}{2}u_{xx} + \mu u_x + g(t, x, u, \sigma u_x) = 0.$$

For, applying Itô's formula to $u(t, X_t)$ we obtain

$$\begin{aligned} du(t, X_t) &= \left(u_t + \frac{\sigma^2}{2}u_{xx} + \mu u_x\right) dt + \sigma u_x dW_t \\ &= -g(t, X_t, u(t, X_t), \sigma u_x(t, X_t))dt + \sigma u_x dW_t \end{aligned}$$

In integral form

$$u(T, X_T) - u(t, X_t) = -\int_t^T g(s, X_s, u(s, X_s), \sigma u_x(s, X_s))ds + \int_t^T \sigma u_x dW_s$$

Taking conditional expectation

$$\mathbb{E}\left[u(T, X_T)|X_t = x\right] - u(t, x) = -\mathbb{E}\left[\int_t^T g(s, X_s, u(s, X_s), \sigma u_x(s, X_s))ds \middle| X_t = x\right]$$

In other words,

$$u(t, x) = \mathbb{E}\left[\varphi(X_T) + \int_t^T g(s, X_s, u(s, X_s), \sigma u_x(s, X_s))ds \middle| X_t = x\right]$$