Lecture 24 – 2020.07.14 – 12:15 via Zoom

Backward SDEs and non-linear PDEs (continued)

Recall notations from the previous lecture

$$\mathscr{L}_t f(t,x) = \sum_{i=1}^{d} b^i(t,x)\nabla^i f(t,x) + \sum_{i,j=1}^{d} a^{i,j}(t,x)\cdot\nabla^i\nabla^j f(t,x), \qquad t\geqslant 0, x\in\mathbb{R}^d,$$

where $f\in C^{1,2}(\mathbb{R}_+\times\mathbb{R}^d;\mathbb{R})$ and $b:\mathbb{R}_+\times\mathbb{R}^d\to\mathbb{R}^d$, $a:\mathbb{R}_+\times\mathbb{R}^d\to\mathbb{R}^{d\times d}$ and b,a are sufficiently regular and $a=\frac{1}{2}\sigma\sigma^T$ for some $\sigma:\mathbb{R}_+\times\mathbb{R}^d\to\mathbb{R}^{d\times d}$.

We consider here a special kind of PDEs, of the form

$$\partial_t u(t,x) + \mathscr{L}u(t,x) + f(t,x,u(t,x),\sigma(t,x)\nabla u(t,x)) = 0 \tag{1}$$

where $\nabla = \mathrm{D}_x$ is the derivative with respect to the space variable (i.e. the gradient).

We argued that if $(X_s^{t,x})_{s\geqslant t}$ is the solution to

$$\mathrm{d}X_s^{t,x} = b(s,X_s^{t,x})\mathrm{d}s + \sigma(s,X_s^{t,x})\mathrm{d}W_s, \qquad s\geqslant t, \tag{2}$$

with

$$X_t^{t,x} = x\in\mathbb{R}^d$$

and if we let $Y_s = u(s,X_s^{t,x})$, $Z_s=\sigma(X^{t,x})\nabla u(t,X_s^{t,x})$ for $s\geqslant t$ the the pair (Y,Z) satisfies the BSDE:

$$\mathrm{d}Y_s = -f(s,X_s^{t,x},Y_s,Z_s)\mathrm{d}s + Z_s\mathrm{d}W_s. \tag{3}$$

This was our motivation to look into the solution theory of a more general class of BSDEs of the form

$$-\mathrm{d}Y_s = f(s,\omega,Y_s,Z_s)\mathrm{d}s - Z_s\mathrm{d}W_s, \qquad Y_T=\xi \tag{4}$$

where $(\Omega,\mathscr{F},\mathbb{P})$ is the canonical d-dimensional Wiener space, $\xi\in L^2(\Omega,\mathscr{F}_T,\mathbb{P};\mathbb{R}^n)=L^2(\mathscr{F}_T;\mathbb{R}^n)$ (i.e. ξ takes values in $\mathbb{R}^n$ and is $\mathscr{F}_T$ measurable) and Y,Z are adapted processes taking values respectively in $\mathbb{R}^n$ and $\mathbb{R}^{n\times d}\approx L(\mathbb{R}^d,\mathbb{R}^n)$. Morever $f:\mathbb{R}_+\times\Omega\times\mathbb{R}^n\times\mathbb{R}^{n\times d}\to\mathbb{R}^n$ (called the *generator* or *driver*) is an adapted process, i.e. $(y,z)\mapsto f(t,\omega,y,z)$ is measurable wrt. $\mathscr{F}_t$. Standard conditions are that

$$f(\cdot,\cdot,0,0)\in L^2_{\mathscr{P}}([0,T]\times\Omega;\mathbb{R}^n) \tag{5}$$

and there exists a constant L such that (Lipshitz condition)

$$|f(t,\omega,y_1,z_1)-f(t,\omega,y_2,z_2)|\leqslant L(|y_1-y_2|+|z_1-z_2|), \qquad y_1,y_2\in\mathbb{R}^n, z_1,z_2\in\mathbb{R}^{n\times d}$$

for almost every (t,ω).

And proved a theorem guarateeing that under these conditions the BSDE (4) has a unique solution

$$(Y,Z)\in L^2_T(\mathbb{R}^n)\times L^2_T(\mathbb{R}^{n\times d}).$$

Representation formula for non-linear PDEs.

We let $(X_s^{t,x})_{s\geqslant 0}$ solving the (forward) SDE

$$\mathrm{d}X_s^{t,x} = b(s,X_s^{t,x})\mathrm{d}s + \sigma(s,X_s^{t,x})\mathrm{d}W_s, \qquad s\geqslant t, \tag{6}$$

for $s \geqslant t$ and such that $X_s^{t,x} = x$ for $s \leqslant t$. For given

$$f : \mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}^n \times \mathbb{R}^{n \times d} \to \mathbb{R}^n$$

and

$$\Psi : \mathbb{R}^d \to \mathbb{R}^n,$$

let $(Y_s^{t,x}, Z_s^{t,x})_{s \in [0,T]}$ the solution of the BSDE ($s \in [0,T]$)

$$-\mathrm{d}Y_s^{t,x} = f(s, X_s^{t,x}, Y_s^{t,x}, Z_s^{t,x})\mathrm{d}s - Z_s^{t,x}\mathrm{d}W_s, \qquad Y_T = \Psi(X_T^{t,x}) \tag{7}$$

This system of a forward SDE and a BSDE is called a (decoupled) forward-backward-SDE (FBSDE), is decoupled because the forward process $(X_s^{t,x})_s$ does not depends on $(Y^{t,x}, Z^{t,x})$ (otherwise is called fully-coupled).

We will assume that σ, b are *Lipshitz and of linear growth*, that f depends in a Lipschitz way on Y, Z (like in the general theory of the previous lecture) and moreover we have that

$$|f(t, x, 0, 0)| + |\Psi(x)| \leqslant C(1 + |x|^p),$$

for some $p \geqslant 1/2$. In this case the generator $f(t, X^{t,x}(\omega), y, z)$ satisfies the condition (5) and the final condition $\Psi(X_T^{t,x})$ is in L^2 because from the general theory of SDEs we can prove that solutions to (6) satisfy

$$\sup_{s \in [0,T]} \mathbb{E}[|X_s^{t,x}|^{2p}] \leqslant K(1 + |x|^{2p})$$

for some $K > 0$. This can be proven easily from a combination of BDG inequality (remember these are the L^p for the stochastic integral) and Grownwall's lemma, via the integral formulation of the SDE exploiting the linear growth of the coefficients b, σ.

From these assumptions it follows that the data of the BSDE satisfy the standard assumptions (those we introduced the last lecture) and therefore by the Theorem we proved it has a unique solution $(Y_s^{t,x}, Z_s^{t,x})_{s \in [0,T]}$.

Observe also that the process $(X_s^{t,x})_{s \in [0,T]}$ is a Markov process (exercise, it follows from the uniqueness of solutions to the SDE) and one has for all $t \leqslant u$

$$X_s^{t, X_t^{u,x}} = X_s^{t,x}, \qquad u \leqslant s$$

almost surely.

We want to prove now that we can express $Y_s^{t,x}, Z_s^{t,x}$ as deterministic functions of $X_s^{t,x}$. Namely that there exists two functions u, v such that $Y_s^{t,x} = u(s, X_s^{t,x})$ and $Z_s^{t,x} = \sigma(s, X_s^{t,x}) v(s, X_s^{t,x})$.

Introduce $(\mathscr{F}_{t,s})_{s \geqslant t}$ to be the completed right-continuous filtration generated by $(W_u - W_t)_{u \geqslant t}$, i.e. the future filtration of W after time t.

Proposition 1. *The solution $(Y_s^{t,x}, Z_s^{t,x})_{s \in [0,T]}$ is $(\mathscr{F}_{t,s})_{s \in [t,T]}$ adapted. In particular $\mathscr{F}_{t,s}$ is $\mathscr{F}_{t,t}$ measureable and therefore deterministic and $(Y_s^{t,x})_{s \in [0,t]}$ is also deterministic.*

Proof. Consider the new Brownian motion $\tilde{W}_s = W_{t+s} - W_t$ and let $\tilde{\mathscr{F}}$ its complected right-contrinuous filtration. Let (X', Y', Z') be the solution to the FBSDE:

$$\mathrm{d}X_s' = b(t + s, X_s')\mathrm{d}s + \sigma(t + s, X_s')\mathrm{d}W_s', \qquad s \geqslant 0, \qquad X_0' = x,$$

$$-\mathrm{d}Y_s' = f(t + s, X_s', Y_s', Z_s')\mathrm{d}s - Z_s'\mathrm{d}W_s, \qquad s \geqslant 0, \qquad Y_{T-t}' = \Psi(X_{T-t}').$$

By the general theory this FBSDE has a unique solution and then it is clear that $X_s' = X_{t+s}^{t,x}$ for $s \in [0, T-t]$ and similarly $(Y_s', Z_s') = (Y_{t+s}^{t,x}, Z_{t+s}^{t,x})$ for $s \in [0, T-t]$. However X', Y', Z' are adapted to $(\tilde{\mathscr{F}}_s)_{s \geqslant 0}$ which means that $(X_{t+s}^{t,x}, Y_{t+s}^{t,x}, Z_{t+s}^{t,x})_{s \geqslant 0}$ is adapted to $(\tilde{\mathscr{F}}_s)_{s \geqslant 0}$ and therefore $(X_s^{t,x}, Y_s^{t,x}, Z_s^{t,x})_{s \in [t,T]}$ is adapted to $(\mathscr{F}_{t,s})_{s \in [s,T]}$ and therefore $(X_t^{t,x}, Y_t^{t,x}, Z_t^{t,x})$ is deterministic.

When $t' \leqslant t$ to see that $(Y_{t'}^{t,x}, Z_{t'}^{t,x})$ is deterministic one can just take $\tilde{W}_s = W_{t'+s} - W_{t'}$ and repeat the above argument by replacing there t with t'. Indeed the crucial remark is that $X_{t'}^{t,x} = x$ for any $t' \leqslant t$. □

Proposition 2. *There exists two deterministic measurable functions u, v such that $Y_s^{t,x} = u(s, X_s^{t,x})$ and $Z_s^{t,x} = \sigma(s, X_s^{t,x}) v(s, X_s^{t,x})$*

Proof. By induction, as follows. Assume first f does not depends on y, z. Then in this case

$$Y_s^{t,x} = \mathbb{E}\left[\int_s^T f(r, X_r^{t,x}) \mathrm{d}r + \Psi(X_T^{t,x}) \middle| \mathscr{F}_s\right] = \mathbb{E}\left[\int_s^T f(r, X_r^{t,x}) \mathrm{d}r + \Psi(X_T^{t,x}) \middle| X_s^{t,x}\right] = u(s, X_s^{t,x})$$

because $(X_s^{t,x})_{s \geqslant 0}$ is a Markov process and we can use the Markov property in the 2nd equality and the 3rd equality is just the statement that there exists a measurable function which represents the conditional expectation wrt. $\sigma(X_s^{t,x})$. Similarly one can show that $Z_s^{t,x} = \sigma(s, X_s^{t,x}) v(s, X_s^{t,x})$. (See Perkowski).

In the general case we introduce an iterative procedure. Define $Y^{(0)} = Z^{(0)} = 0$ then define $(Y^{(k+1)}, Z^{(k+1)})$ and the solution of the BSDE with driver $f(r, X_r^{t,x}, Y^{(k)}, Z^{(k)})$. We know from the proof of existence and uniqueness that there exists only one fixed point for this iteration and therefore $(Y^{(k)}, Z^{(k)}) \to (Y^{t,x}, Z^{t,x})$ (if you want this is the Picard iteration to construct the solution to the BSDE). From this we deduce that there exists functions u_k, v_k such that $Y_s^{(k)} = u_k(s, X_s^{t,x})$ and $Z_s^{(k)} = \sigma(s, X_s^{t,x}) v_k(s, X_s^{t,x})$, and the is not difficult to pass to the limit by letting $u^i(s, x) := \limsup_{k \to \infty} (u_k(s, x))^i$ (componentwise) and then $u^i(s, X_s^{t,x}) = \lim_{k \to \infty} Y_s^{(k)} = Y_s^{t,x}$ by convergence of the Picard iterations. Similarly one reason for the sequence $Z^{(k)}$ to deduce that

$$Z_s^{t,x} = \lim_{k \to \infty} Z_s^{(k)} = \sigma(s, X_s^{t,x}) \lim_{k \to \infty} v_k(s, X_s^{t,x}) = \sigma(s, X_s^{t,x})\, v(s, X_s^{t,x}).$$

This concludes the proof. □

Finally it remains to identify the functions u, v as associated to a nonlinear PDE.

We reason as follows: let u be the solution of the semilinar parabolic PDE

$$\partial_t u(t,x) + \mathscr{L}_t u(t,x) + f(t, x, u(t,x), \sigma(t,x) \nabla u(t,x)) = 0, \qquad t \in [0,T], x \in \mathbb{R}^d$$

with *final* condition $u(T, x) = \Psi(x)$.

Theorem 3. *(Generalised Feynman-Kac formula for quasilinear equations) Assume that $u \in C^{1,2}([0,T] \times \mathbb{R}^d; \mathbb{R}^n)$ is a solution to the PDE (2) such that*

$$|u(s,x)| + |\sigma(s,x) \nabla u(s,x)| \leqslant C(1 + |x|^k)$$

for some $k \geqslant 1$. Then if $(X_s^{t,x}, Y_s^{t,x}, Z_s^{t,x})_{s \in [0,T]}$ is the unique solution to the FBSDE with final condition Ψ and driver f then we have

$$Y_s^{t,x} = u(s, X_s^{t,x}), \qquad Z_s^{t,x} = \sigma(s, X_s^{t,x}) \nabla u(s, X_s^{t,x}), \qquad s, t \in [0,T], x \in \mathbb{R}^d.$$

In particular

$$u(t,x) = Y_t^{t,x}, \qquad t \in [0,T], x \in \mathbb{R}^d,$$

and therefore the PDE has a unique solution.

Proof. We apply Ito formula

$$\mathrm{d}u(s, X_s^{t,x}) = (\partial_s + \mathscr{L}_s) u(s, X_s^{t,x}) \mathrm{d}s + \sigma(s, X_s^{t,x}) \nabla u(s, X_s^{t,x}) \mathrm{d}W_s$$

$$= -f(s, X_s^{t,x}, u(s, X_s^{t,x}), \sigma(s, X_s^{t,x}) \nabla u(s, X_s^{t,x})) \mathrm{d}s + \sigma(s, X_s^{t,x}) \nabla u(s, X_s^{t,x}) \mathrm{d}W_s$$

which means that the pair $(u(s, X_s^{t,x}), \sigma(s, X_s^{t,x}) \nabla u(s, X_s^{t,x}))$ is a solution to the BSDE, the final condition is ok since $u(T, X_T^{t,x}) = \Psi(X_T^{t,x})$ and by uniqueness we have $(u(s, X_s^{t,x}), \sigma(s, X_s^{t,x}) \nabla u(s, X_s^{t,x})) = (Y_s^{t,x}, Z_s^{t,x})$ for all $s \in [0,T]$. □

Remark 4. With stronger conditions on the coefficients of the PDE one can prove directly that given a solution to the BSDE which then, as we have seen can always be represented as $Y_s^{t,x} = u(s, X_s^{t,x})$ and $Z_s^{t,x} = \sigma(s, X_s^{t,x}) v(s, X_s^{t,x})$ for *some* functions u, v, then one necessarily have that $u \in C^{1,2}$ and $v = \nabla u$ and u solves the PDE. (see the notes of Perkowski for some literature on this).

(the following is not in the exam)

Rough path theory

Rough path theory is a way to make sense of SDEs without using stochastic integrals.

Imagine you want to give an "analytic" meaning to the equation (let's ignore the drift b)

$$\mathrm{d}X_t = \sigma(X_t)\mathrm{d}W_t, \qquad X_0 = x,$$

where W is a Brownian motion or possibly a similar process which is nowhere differentiable and maybe not even a semimartingale.

Recall that stochastic integrals are only defined almost surely (or a limit in probability).

- Extend SDE theory beyond the semimartingale setting
- Have a robust theory of SDEs (meaning that I can reliably approximate a stochastic integral)
- Prove Wong-Zakai type theorems, i.e. let $W^\varepsilon \to W$ (as $\varepsilon \to 0$) to be smooth approximations of Brownian motion and let X^ε be the solution of the ODE

  $$\partial_t X_t^\varepsilon = \sigma(X_t) \partial_t W_t^\varepsilon, \qquad X_0 = x.$$

  Then we want to prove that $X^\varepsilon \to X$ where X solve the SDE above. In general this is false!!.

For example Wong-Zakai ('70) proved that if

$$W_t^\varepsilon = \int \varepsilon^{-1} \rho((t-s)/\varepsilon) W_s \mathrm{d}s$$

where $\rho: \mathbb{R} \to \mathbb{R}_+$, smooth and with integral one. Then $W_t^\varepsilon \to W_t$ as $\varepsilon \to 0$ for all t almost surely (and actually almost sure convergence takes place in any Hölder space with index less that $1/2$), but nonetheless one as that $X^\varepsilon \to Y$ where Y is the process which solves the SDE

$$\mathrm{d}Y_t^i = \sum_{\alpha=1}^{n} \sigma_\alpha^i(Y_t)\mathrm{d}W_t^\alpha + \frac{1}{2} C_\rho \sum_{\alpha=1}^{n} \sum_{j=1}^{d} \sigma_\alpha^j(Y_t) \nabla^j \sigma_\alpha^i(Y_t)\mathrm{d}t, \qquad t \geqslant 0, i = 1, \ldots, d$$

where here I'm assuming that W takes values in $\mathbb{R}^n$ and Y in $\mathbb{R}^d$ and $\sigma_\alpha: \mathbb{R}^d \to \mathbb{R}^d$ for $\alpha = 1, \ldots, n$ smooth. The constant C_ρ depends on ρ.
