Backward SDEs and non-linear PDEs (continued)

Recall notations from the previous lecture

\[ \mathcal{L} f(t, x) = \sum_{i=1}^{d} b_i(t, x) \nabla f(t, x) + \sum_{i,j=1}^{d} a_{ij}(t, x) \cdot \nabla^2 f(t, x), \quad t \geq 0, x \in \mathbb{R}^d, \]

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

\[ \frac{\partial u(t, x)}{\partial t} + \mathcal{L} u(t, x) + f(t, x, u(t, x), \sigma(t, x) \nabla u(t, x)) = 0 \quad \text{(1)} \]

where \( \nabla = D_x \) is the derivative with respect to the space variable (i.e. the gradient).

We argued that if \( (X_t^{t,x})_{t \geq 0} \) is the solution to

\[ dX_t^{t,x} = b(s, X_t^{t,x})ds + \sigma(s, X_t^{t,x})dW_s, \quad s \geq t, \quad (2) \]

with

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

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

\[ dY_s = -f(s, X_s^{s,x}, Y_s, Z_s)dW_s + Z_s dW_s \quad \text{(3)} \]

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

\[ -dY_s = f(s, \omega, Y_s, Z_s)ds - Z_s dW_s, \quad Y_T = \xi \quad \text{(4)} \]

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

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

and there exists a constant \( L \) such that (Lipschitz condition)

\[ |f(t, \omega, y_1, z_1) - f(t, \omega, y_2, z_2)| \leq L(|y_1 - y_2| + |z_1 - z_2|), \quad y_1, y_2 \in \mathbb{R}^n, z_1, z_2 \in \mathbb{R}^{n \times d} \]

for almost every \((t, \omega)\).

And proved a theorem guaranteeing that under these conditions the BSDE (4) has a unique solution \((Y, Z) \in L^2(\mathbb{R}^n) \times L^2(\mathbb{R}^{n \times d})\).

Representation formula for non-linear PDEs.

We let \((X_t^{t,x})_{t \geq 0}\) solving the (forward) SDE

\[ dX_t^{t,x} = b(s, X_t^{t,x})ds + \sigma(s, X_t^{t,x})dW_s, \quad s \geq t, \quad \text{(6)} \]
for $s \geq t$ and such that $X_{s}^{t,x} = x$ for $s \leq t$. 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 $(Y^{t,x}_{s}, Z^{t,x}_{s})_{s \in [0,T]}$ the solution of the BSDE $(s \in [0,T])$

$$-dY^{t,x}_{s} = f(s, X^{t,x}_{s}, Y^{t,x}_{s}, Z^{t,x}_{s})ds - Z^{t,x}_{s}dW_{s}, \quad Y_{T} = \Psi(X^{t,x}_{T}) \quad (7)$$

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

We will assume that $\sigma, b$ are Lipschitz and of linear growth, that $f$ depends in a Lipschitz way on $Y, Z$ (like in the general theory of the procedure lecture) and moreover we have that

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

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

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

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 Gronwall's lemma, via the integral formulation of the SDE exploiting the linear growth of the coefficients $b, \sigma$.

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 $(Y^{t,x}_{s}, Z^{t,x}_{s})_{s \in [0,T]}$.

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

$$X^{t,x}_{s} = X^{t,x}_{u}, \quad u \leq s$$

almost surely.

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

Introduce $(\mathcal{F}_{t,s})_{s \geq t}$ to be the completed right-continuous filtration generated by $(W_{u} - W_{t})_{u \geq t}$, i.e. the future filtration of $W_{t}$ after time $t$.

**Proposition 1.** The solution $(Y^{t,x}_{s}, Z^{t,x}_{s})_{s \in [0,T]}$ is $(\mathcal{F}_{t,s})_{s \in [t,T]}$ adapted. In particular $\mathcal{F}_{t,s}$ is $\mathcal{F}_{t,\tilde{s}}$ measurable and therefore deterministic and $(Y^{t,x}_{s})_{s \in [0,T]}$ is also deterministic.

**Proof.** Consider the new Brownian motion $\tilde{W}_{s} = W_{s+t} - W_{t}$ and let $\tilde{\mathcal{F}}$ its completed right-continuous filtration. Let $(X^{t,x}_{s}, Y^{t,x}_{s}, Z^{t,x}_{s})$ be the solution to the FBSDE:

$$dX^{t}_{s} = b(t+s, X^{t}_{s})ds + \sigma(t+s, X^{t}_{s})dW^{t}_{s}, \quad s \geq 0, \quad X^{t}_{0} = x,$$

$$-dY^{t}_{s} = f(t+s, X^{t}_{s}, Y^{t}_{s}, Z^{t}_{s})ds - Z^{t}_{s}dW^{t}_{s}, \quad s \geq 0, \quad Y^{t}_{T} = \Psi(X^{t,x}_{T}) .$$

By the general theory this FBSDE has a unique solution and then it is clear that $X^{t}_{s} = X^{t,x}_{s}$ for $s \in [0, T-t]$ and similarly $(Y^{t}_{s}, Z^{t}_{s}) = (Y^{t,x}_{s}, Z^{t,x}_{s})$ for $s \in [0, T-t]$. However $X_{s}, Y_{s}, Z_{s}$ are adapted to $(\mathcal{F}_{t,s})_{s \geq 0}$ which means that $(X^{t,x}_{s}, Y^{t,x}_{s}, Z^{t,x}_{s})_{s \geq 0}$ is adapted to $(\mathcal{F}_{t,s})_{s \geq 0}$ and therefore $(X^{t,x}_{s}, Y^{t,x}_{s}, Z^{t,x}_{s})_{s \in [t,T]}$ is adapted to $(\mathcal{F}_{t,s})_{s \in [t,T]}$ and therefore $(X^{t,x}_{s}, Y^{t,x}_{s}, Z^{t,x}_{s})_{s \in [t,T]}$ is deterministic.
When $t' \leq t$ to see that $(Y_{t',s}^{i,s}, Z_{t',s}^{i,s})$ is deterministic one can just take $\bar{W}_i = W_{t'-s} - W_i$ and repeat the above argument by replacing there $t'$ with $t'$. Indeed the crucial remark is that $X_{t',s}^{i,s} = x$ for any $t' \leq t$.

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

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

$$Y_t^{t,s} = \mathbb{E}\left[ \int_s^T f(r, X_r^{t,s}) dr + \Psi(X_T^{t,s}) \bigg| F_s \right] = \mathbb{E}\left[ \int_s^T f(r, X_r^{t,s}) dr + \Psi(X_T^{t,s}) \right] = u(s, X_s^{t,s})$$

because $(X_t^{t,s})_{t \geq 0}$ 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. $\sigma(X_s^{t,s})$. Similarly one can show that $Z_t^{t,s} = \sigma(s, X_t^{t,s})v(s, X_t^{t,s})$. (See Perkowski).

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

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

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 semilinear parabolic PDE

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

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

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

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

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

In particular

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

and therefore the PDE has a unique solution.

Proof. We apply Itô formula

$$du(s, X_s^{t,s}) = (\partial_s + \mathcal{L}_i)u(s, X_s^{t,s}) ds + \sigma(s, X_s^{t,s}) \nabla u(s, X_s^{t,s}) dW_s$$

$$= -f(s, X_s^{t,s}, u(s, X_s^{t,s}), \sigma(s, X_s^{t,s})u(s, X_s^{t,s})) ds + \sigma(s, X_s^{t,s}) \nabla u(s, X_s^{t,s}) dW_s$$

which means that the pair $(u(s, X_s^{t,s}), \sigma(s, X_s^{t,s})v(s, X_s^{t,s}))$ is a solution to the BSDE, the final condition is ok since $u(T, X_T^{t,s}) = \Psi(X_T^{t,s})$ and by uniqueness we have $(u(s, X_s^{t,s}), \sigma(s, X_s^{t,s})v(s, X_s^{t,s})) = (Y_s^{t,s}, Z_s^{t,s})$ 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 $Y_t^{x,s} = u(s, X_t^{x,s})$ and $Z_t^{x,s} = \sigma(s, X_t^{x,s}) v(s, X_t^{x,s})$ 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).

---

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 $b$)

$$dX_t = \sigma(X_t) dW_t, \quad X_0 = x,$$

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 $W^\epsilon \to W$ (as $\epsilon \to 0$) to be smooth approximations of Brownian motion and let $X^\epsilon$ be the solution of the ODE

$$\partial_t X_t^\epsilon = \sigma(X_t) \partial_t W_t^\epsilon, \quad X_0 = x.$$

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

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

$$W_t^\epsilon = \int e^{-1} \rho((t-s)/\epsilon) W_s ds$$

where $\rho: \mathbb{R} \to \mathbb{R}_+$, smooth and with integral one. Then $W^\epsilon \to W$ as $\epsilon \to 0$ 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 $X^\epsilon \to Y$ where $Y$ is the process which solves the SDE

$$dY^\epsilon_t = \sum_{a=1}^n \sigma^a(Y_t) dW^\epsilon_t + \frac{1}{2} C^2 \sum_{a=1}^n \sum_{j=1}^d \sigma^a_j(Y_t) \nabla_j \sigma^a_i(Y_t) dt, \quad t \geq 0, i = 1, \ldots, d$$

where here I’m assuming that $W$ takes values in $\mathbb{R}^n$ and $Y$ in $\mathbb{R}^d$ and $\sigma^a_i : \mathbb{R}^d \to \mathbb{R}^d$ for $a = 1, \ldots, n$ smooth. The constant $C^2$ depends on $\rho$.