DeepBSDE: A Neural Network-Based Model for Backward Stochastic Differential Equations

YongJae Song
Konkuk University
ORLAB

November 11, 2024

Outline

What is a BSDE?

BSDE Definition

BSDE:

- Defined on a filtered probability space $(\Omega, (\mathcal{F}_t)_{t\in[0,T]}, P)$
- W: d-dimensional Brownian motion

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s)ds - \int_t^T Z_s dW_s, \quad t \in [0, T].$$

Solving a BSDE

Solving

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s)\, ds - \int_t^T Z_s\, dW_s, \quad t \in [0, T]$$

means finding $(Y_t, Z_t)_{t\in[0,T]}$ that satisfies the relation, or equivalently, finding $u(t,x)$ and $v(t,x)$ such that

$$Y_t = u(t, W_t), \quad Z_t = v(t, W_t).$$

Forward BSDE

Forward BSDEs (FBSDEs) generalize BSDEs:
We sought to find $(X_t, Y_t, Z_t)_{t\in[0,T]}$ which satisfy

$$Y_t = \Phi_{X_T} + \int_t^T f(s, X_s, Y_s, Z_s)\, ds - \int_t^T Z_s\, dW_s,$$

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

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

Solving a FBSDE

We want to find l, u, v such that

$$X_t = l(t, W_t), \quad Y_t = u(t, W_t), \quad Z_t = v(t, W_t)$$

or just u, v such that

$$Y_t = u(t, X_t), \quad Z_t = v(t, X_t)$$

since sample paths of X can be easily obtained through forward methods (Euler–Maruyama).

Application: Option Pricing

Consider pricing a European option with

- risk-free interest rate r_t
- risky asset $dS_t = S_t \mu_t dt + S_t \sigma_t dW_t$
- payoff ξ at time T

Let Y be the wealth process of a self-financing portfolio with ϕ_t amount of money invested in S at time t:

$$dY_t = \frac{\phi_t}{S_t} dS_t + r_t (Y_t - \phi_t) dt$$

$$= \left(\phi_t(\mu_t - r_t) + r_t Y_t\right) dt + \phi_t \sigma_t dW_t$$

$$= \left(Z_t \pi_t + r_t Y_t\right) dt + Z_t dW_t$$

where $Z = \phi\sigma$ and assume $\exists$ measurable π with $\mu - r = \sigma\pi$.

Application: Option Pricing

If $Y_T = \xi$, by no-arbitrage principle, the price of the European option is

$$Y_t = \xi - \int_t^T (Z_s \pi_s + r_s)\, ds - \int_t^T Z_t dW_t.$$

Therefore, pricing the European option amounts to solving this BSDE.

Application: Stochastic Control

Let X denote the solution to

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

and consider the Optimization problem $\sup_k J(k)$ where

$$J(k) = \mathbb{E}\left[\Phi(X_T) + \int_0^T f(s, X_s, k_s)\, ds\right].$$

Then $J(k) = Y_0^k$ where Y_t^k is the solution to

$$Y_t^k = \Phi(X_T) + \int_t^T f(s, X_s, k_s)\, ds - \int_t^T Z_s^k\, dB_s.$$

Therefore, the stochastic control problem is related to FBSDEs.

Interpretation of FBSDE

Solving a FBSDE is finding a control process Z and starting point y, so that under the dynamics

$$dY_t = -f(t, X_t, Y_t, Z_t)\,dt + Z_t\,dW_t, \quad dX_t = \mu(t, X_t)\,dt + \sigma(t, X_t)\,dW_t.$$

we have $Y_T^{y,Z} = \Phi(X_T)$ (Y is determined by its starting point y and the control process Z). This can be interpreted as a stochastic optimization problem

$$\inf_{y,Z} \mathbb{E}\left[|\Phi(X_T) - Y_T^{y,Z}|^2\right].$$

Theoretical Solutions

Existence and Uniqueness

Theorem

Assume the following conditions:

- μ, σ, f and Φ are uniformly Lipschitz continuous in (x, y, z),
- $\mu(\cdot, 0), \sigma(\cdot, 0), f(\cdot, 0, 0, 0)$ and $\Phi(0)$ are bounded,
- μ, σ and f are uniformly Hölder-$(1/2)$ continuous in t.

Conclusion:

- There exists a unique solution (Y, Z) to the Forward-Backward Stochastic Differential Equation (FBSDE).

PDE Equivalence

Additionally, if we assume polynomial growth of the coefficients, then solving the BSDE becomes equivalent to solving a parabolic PDE:

Theorem

Let

$$\mathcal{L}_t u(t,x) = \langle \mu(t,x), \nabla u(t,x) \rangle + \frac{1}{2} \mathrm{tr}\left[\sigma^{\otimes 2}(t,x) Hess(u(t,x))\right],$$

and $u(t,x)$ the solution to

$$\begin{cases} \left(\dfrac{\partial}{\partial t} + \mathcal{L}_t\right) u(t,x) + f\left(t, x, u(t,x), (\nabla u(t,x))^\top \sigma(t,x)\right) = 0, \\ u(T,x) = \Phi(x). \end{cases}$$

Then, the solutions to the FBSDE are given by

$$Y(t, X_t) = u(t, X_t), \quad Z_t = \nabla u(t, X_t)^\top \sigma(t, X_t).$$

Numerical Solutions

Euler–Maruyama

To find a numerical solution to forward SDE

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

we discretize

- $0 = t_0 < t_1 < \cdots < t_n = T$

and go forward by sampling:

- Sample $\Delta W_{t_i} := W_{t_i} - W_{t_{i-1}} \sim N(0, t_i - t_{i-1})$
- $X_0 = x$
- $X_{t_i} = X_{t_{i-1}} + \mu(t_{i-1}, X_{t_{i-1}})\Delta t_{i-1} + \sigma(t_{i-1}, X_{t_{i-1}})\Delta W_{t_{i-1}}$

for $i = 1, \ldots, n$. This is easy.

Trouble with BSDE

However, for BSDE

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s)\, ds - \int_t^T Z_s\, dW_s, \quad t \in [0, T]$$

we cannot simply do backwards Euler–Maruyama since integrating backwards may not make sense! Take a look at

$$\int_T^t Z_{-s}\, dW_{-s}.$$

Z_{-s} is not predictable in general, perhaps not even adapted!

Recall that we used left values to define integrals, not right ones.

Trouble with BSDE

Instead, to go backwards, we need to use

$$
\begin{aligned}
Y_{t_{i-1}} &= \mathbb{E}\left[Y_{t_i} + \int_{t_{i-1}}^{t_i} f(s, Y_s, Z_s)\, ds - \int_{t_{i-1}}^{t_i} Z_s\, dW_s \,\middle|\, \mathcal{F}_{t_{i-1}} \right] \\
&= \mathbb{E}\left[Y_{t_i} + \int_{t_{i-1}}^{t_i} f(s, Y_s, Z_s)\, ds \,\middle|\, \mathcal{F}_{t_{i-1}} \right] \\
&\approx \mathbb{E}\left[Y_{t_i} + f(t_{i-1}, Y_{t_i}, Z_{t_{i-1}}) \Delta t_{i-1} \,\middle|\, \mathcal{F}_{t_{i-1}} \right]
\end{aligned}
$$

and similarly for Z.
Calculating coupled conditional expectations is not easy. This is where the obstacle lies:

- Going forward is easy
- Going backward is difficult

Example: Deep FBSDE

First, we choose a discretization of time:

$$0 = t_0 < t_1 < \cdots < t_n = T.$$

Then, we obtain the discrete approximation

$$\{\bar{X}_{t_k}\}_{0 \le k \le n}$$

of the solution to the forward SDE by Euler–Maruyama method.

Example: Deep FBSDE

Next, we construct n neural networks $Z_{t_k}^{\theta_k}$ for approximating the control process at the discrete times by

$$\bar{Z}_{t_k} = Z_{t_k}^{\theta_k}(\bar{X}_{t_k}).$$

To obtain approximations of Y, we execute Euler–Maruyama method on the backward equation:

$$\bar{Y}_{t_0}^{\Theta} = y$$

$$\bar{Y}_{t_{k+1}}^{\Theta} = \bar{Y}_{t_k}^{\Theta} - f(t_k, \bar{X}_{t_k}, \bar{Y}_{t_k}^{\Theta}, Z_{t_k}^{\theta_k}(\bar{X}_{t_k}))(t_{k+1} - t_k) + Z_{t_k}^{\theta_k}(\bar{X}_{t_k})(W_{k+1} - W_k)$$

where $\Theta = (y, \theta_1, \theta_2, \dots, \theta_n)$.

Example: Deep FBSDE

Finally, we train the neural networks to solve the minimization problem.

$$\inf_{y,Z} \mathbb{E}[|\Phi(X_T) - Y_T^{y,Z}|^2] \approx \inf_{\Theta} \mathbb{E}[|\Phi(\bar{X}_T) - \bar{Y}_T^{\Theta}|^2].$$

Empirically, we solve

$$\inf_{\Theta} \frac{1}{M} \sum_{l=1}^{M} \mathbb{E}[|\Phi(\bar{X}_{T,l}) - \bar{Y}_{T,l}^{\Theta}|^2],$$

where $\bar{X}_{T,l}, \bar{Y}_{T,l}^{\Theta}\,(l = 1, \dots, M)$ are obtained through sampling $\bar{X}_T$ and $\bar{Y}_T^{\Theta}$ M times.

Summary

Key Concepts:

- **BSDEs:** Backward stochastic differential equations link a terminal condition to dynamics driven by Brownian motion.
- **FBSDEs:** Forward-backward systems generalize BSDEs, combining forward SDEs with backward dynamics.
- **Applications:** BSDEs are fundamental in option pricing, stochastic control, and optimization problems.

Theoretical Insights:

- Existence and uniqueness rely on Lipschitz and boundedness conditions.
- Equivalence with parabolic PDEs provides a bridge to classical methods.

Numerical Challenges and Solutions:

- Forward SDEs are straightforward with Euler–Maruyama methods.
- BSDEs require backward integration, making conditional expectations critical.
- Neural networks in DeepBSDE offer an efficient framework for solving these equations.