Backward Stochastic Volterra Integral Equations

Jiongmin Yong

University of Central Florida

July, 2010

Outline

1. Introduction — Motivations
2. Definition of Solution
3. Well-Posedness of BSVIEs
4. Properties of Solutions
5. Some Remarks

1. Introduction — Motivations

$(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ — a complete filtered probability space
$W(\cdot)$ — a one-dimensional standard Brownian motion
$\mathbb{F} \equiv \{\mathcal{F}_t\}_{t\geq 0}$ — natural filtration of $W(\cdot)$, augmented by all $\mathbb{P}$-null sets

Consider FSDE:

$$(1.1) \qquad \begin{cases} dX(t) = b(t, X(t))dt + \sigma(t, X(t))dW(t), \\ X(0) = x. \end{cases}$$

Equivalent to:

$$(1.2) \qquad X(t) = x + \int_0^t b(s, X(s))ds + \int_0^t \sigma(s, X(s))dW(s).$$

General forward stochastic Volterra integral equation: (FSVIE)

$$(1.3) \quad X(t) = \varphi(t) + \int_0^t b(t, s, X(s))ds + \int_0^t \sigma(t, s, X(s))dW(s).$$

- In general, FSVIE (1.3) cannot be transformed into a form of FSDE (1.1).
- FSVIE (1.3) allows some long-range dependence on the noises.
- Could allow $\sigma(t, s, X(s))$ to be $\mathcal{F}_t$-measurable, still might have adapted solutions (Pardoux–Protter, 1990).
- May model wealth process involving investment delay, etc. (Duffie–Huang, 1986).

Consider BSDE:

$$(1.4)\quad \begin{cases} dY(t) = -g(t, Y(t), Z(t))dt + Z(t)dW(t), & t \in [0, T], \\ Y(T) = \xi. \end{cases}$$

- Linear case was introduced by Bismut (1973).
- Nonlinear case was introduced by Pardoux–Peng (1990).
- Can be applied to (European) contingent claim pricing, stochastic differential utility, dynamic risk measures,...
- Leads to nonlinear Feynman-Kac formula, pointwise convergence in homogenization problems, nonlinear expectation, ...

BSDE (1.4) is equivalent to

$$(1.5) \qquad Y(t) = \xi + \int_t^T g(s, Y(s), Z(s))ds - \int_t^T Z(s)dW(s).$$

Called a backward stochastic Volterra integral equation (BSVIE).

Recall:

$$(1.2) \qquad X(t) = x + \int_0^t b(s, X(s))ds + \int_0^t \sigma(s,X(s))dW(s).$$

$$(1.3) \quad X(t) = \varphi(t) + \int_0^t b(t, s, X(s))ds + \int_0^t \sigma(t, s, X(s))dW(s).$$

Question:

What is the analog of (1.3) for (1.5) as (1.3) for (1.2)?

A Proposed Form:

$$Y(t) = \psi(t) + \int_t^T g(t, s, Y(s), Z(t, s), Z(s, t))ds - \int_t^T Z(t, s)dW(s), \quad t \in [0, T], \tag{1.6}$$

$(Y(\cdot), Z(\cdot\,, \cdot))$ — unknown process

Remarks:

- The term $Z(t, s)$ depends on t and s;
- The drift depends on both $Z(t, s)$ and $Z(s, t)$.
- (1.6) is strictly more general than BSDE (1.5).
- $\psi(\cdot)$ does not have to be $\mathbb{F}$-adapted.
- Need $Z(t, \cdot)$ to be $\mathbb{F}$-adapted, and $\int_0^T |Z(t, s)|^2 ds < \infty$, a.e. $t \in [0, T]$, a.s.

By taking conditional expectation on (1.6), we have

$$Y(t) = \mathbb{E}\Big[\psi(t) + \int_t^T g(t,s,Y(s),Z(t,s),Z(s,t))ds\Big|\mathcal{F}_t\Big].$$

This leads to the **second** interesting motivation.

- Expected discounted utility (process) has the form:

$$Y(t) = \mathbb{E}\Big[\xi e^{-\beta(T-t)} + \int_t^T u(C(s))e^{-\beta(s-t)}ds\Big|\mathcal{F}_t\Big], \quad t \in [0,T].$$

$C(\cdot)$ — consumption process, $u(\cdot)$ — utility function
β — discount rate, ξ — terminal time wealth

- Expected discounted utility is equivalent to a linear BSDE:

$$Y(t) = \xi + \int_t^T \big[-\beta Y(s) + C(u(s))\big]ds - \int_t^T Z(s)dW(s).$$

- $e^{-\beta(s-t)}$ exhibits a time-consistent memory effect. If the memory is not time-consistent, the utility process will not be a solution to a BSDE! But, it might be a solution to a BSVIE!
- Duffie–Epstein (1992) introduced stochastic differential utility:

$$Y(t) = \mathbb{E}\Big[\xi + \int_t^T g(s, Y(s))ds \Big| \mathcal{F}_t\Big], \quad t \in [0, T].$$

which is equivalent to a nonlinear BSDE:

$$Y(t) = \xi + \int_t^T g(s, Y(s))ds - \int_t^T Z(s)dW(s).$$

The Third Motivation:

Consider controlled FSVIE:

$$X(t)=\varphi(t)+\int_0^t b(t,s,X(s),u(s))ds \\ +\int_0^t \sigma(t,s,X(s),u(s))dW(s).$$

To state the first order necessary condition (of Pontryagin type) for the corresponding optimal control problem, one needs the *adjoint equation* which should be a BSVIE.

2. Definition of Solutions.

Let $H = \mathbb{R}^m, \mathbb{R}^{m\times d}$, etc., with norm $|\cdot|$.

$$L^2(\Omega) = \big\{\xi : \Omega \to H \;\big|\; \xi \in \mathcal{F}_T, E|\xi|^2 < \infty\big\},$$

$$\begin{aligned} L^2((0,T)\times\Omega) = \big\{\varphi : (0,T)\times\Omega \to H \;\big|\; \\ \varphi \text{ is } \mathcal{B}([0,T])\otimes\mathcal{F}_T\text{-measurable, } \mathbb{E}\int_0^T |\varphi(t)|^2 dt < \infty\big\}, \end{aligned}$$

$$L^2_{\mathbb{F}}(0,T) = \big\{\varphi \in L^2((0,T)\times\Omega), \varphi(\cdot) \text{ is } \mathbb{F}\text{-adapted}\big\}.$$

$$\begin{aligned} L^2(0,T;L^2_{\mathbb{F}}(0,T)) = \big\{Z; [0,T]^2\times\Omega \to H \;\big|\; \\ Z(t,\cdot) \text{ is } \mathbb{F}\text{-adapted, a.e. } t\in[0,T], \\ \mathbb{E}\int_0^T\int_0^T |Z(t,s)|^2 ds dt < \infty\big\}. \end{aligned}$$

Recall:

$$
\begin{aligned}
Y(t) = \psi(t) + \int_t^T g(t,s,Y(s),Z(t,s),Z(s,t))ds \\
- \int_t^T Z(t,s)dW(s), \qquad t \in [0,T],
\end{aligned}
\tag{2.1}
$$

Similar to BSDEs, it seems to be reasonable to introduce

Definition 2.1. $(Y,Z) \in L^2_{\mathbb{F}}(0,T) \times L^2(0,T;L^2_{\mathbb{F}}(0,T))$ satisfying (2.1) is called an *adapted solution* of BSVIE (2.1).

Example 2.2. Consider BSVIE:

$$Y(t) = \int_t^T Z(s,t)ds - \int_t^T Z(t,s)dW(s), \qquad t \in [0,T]. \tag{2.2}$$

We can check that

$$\begin{cases} Y(t) = (T-t)\zeta(t), & t \in [0,T], \\ Z(t,s) = I_{[0,t]}(s)\zeta(s), & (t,s) \in [0,T] \times [0,T], \end{cases}$$

is an adapted solution of (2.2) for any $\zeta(\cdot) \in L^2_{\mathbb{F}}(0,T;\mathbb{R})$. Thus, adapted solutions are not unique!

Observation:

$$Y(t) = \psi(t) + \int_t^T g(t, s, Y(s), Z(t, s), Z(s, t))ds \tag{2.1}$$
$$- \int_t^T Z(t, s)dW(s), \quad t \in [0, T],$$

does not give enough "restrictions" on $Z(t, s)$ with $0 \le s \le t \le T$.

Need to "specify" $Z(t, s)$ for $0 \le s \le t \le T$.

Definition 2.3. $(Y, Z) \in L^2_{\mathbb{F}}(0, T) \times L^2(0, T; L^2_{\mathbb{F}}(0, T))$ is called an *adapted M-solution* of (2.1) if (2.1) is satisfied and also

$$Y(t) = \mathbb{E}Y(t) + \int_0^t Z(t, s)dW(s), \qquad t \in [0, T]. \tag{2.3}$$

3. Well-posedness of BSVIEs.

(H1) Map g is measurable satisfying

$$\mathbb{E}\int_0^T \Big(\int_t^T |g(t,s,0,0)|ds\Big)^2 dt < \infty,$$

and exists a (deterministic) function L with

$$\sup_{t\in[0,T]}\int_t^T L(t,s)^{2+\varepsilon}ds < \infty,$$

for some $\varepsilon > 0$ such that

$$\begin{aligned}&|g(t,s,y,z,\zeta) - g(t,s,\bar{y},\bar{z},\bar{\zeta})|\\ &\le L(t,s)\big(|y-\bar{y}| + |z-\bar{z}| + |\zeta-\bar{\zeta}|\big).\end{aligned}$$

Theorem 3.1. Let (H1) hold. Then $\forall\psi$, (2.1) admits a unique adapted M-solution (Y, Z). Moreover: for any $r \in [0, T]$,

$$
\begin{aligned}
&\int_r^T \mathbb{E}|Y(t)|^2 dt + \int_r^T \int_r^T \mathbb{E}|Z(t,s)|^2 ds dt \\
&\le C\Big[\int_r^T \mathbb{E}|\psi(t)|^2 dt + \int_r^T \Big(\int_r^T |g(t,s,0,0)| ds\Big)^2 dt\Big].
\end{aligned} \tag{3.1}
$$

If $(\bar{Y}, \bar{Z})$ is the adapted M-solution corresponding to $\bar{\psi}$, then

$$
\begin{aligned}
&\int_r^T \mathbb{E}|Y(t) - \bar{Y}(t)|^2 dt + \int_r^T \int_r^T \mathbb{E}|Z(t,s) - \bar{Z}(t,s)|^2 ds dt \\
&\le C \int_r^T \mathbb{E}|\psi(t) - \bar{\psi}(t)|^2 dt, \qquad \forall r \in [0, T].
\end{aligned} \tag{3.2}
$$

A Difference between BSDEs and BSVIEs:

For BSDE

$$\begin{aligned}
Y(t) &= \xi + \int_t^T g(s, Y(s), Z(s))ds - \int_t^T Z(s)dW(s) \\
&= \xi + \int_{T-\delta}^T g(s, Y(s), Z(s))ds - \int_{T-\delta}^T Z(s)dW(s) \\
&\quad + \int_t^{T-\delta} g(s, Y(s), Z(s))ds - \int_t^{T-\delta} Z(s)dW(s) \\
&= Y(T-\delta) + \int_t^{T-\delta} g(s, Y(s), Z(s))ds - \int_t^{T-\delta} Z(s)dW(s), \\
&\qquad t \in [0, T-\delta].
\end{aligned}$$

Thus, one can obtain the solvability on $[T-\delta, T]$, then on $[T-2\delta, T-\delta]$, etc., to get solvability on $[0, T]$.

For BSVIE: (with $t \in [0, T-\delta]$)

$$
\begin{aligned}
Y(t) &= \psi(t)+\int_t^T g(t,s,Y(s),Z(t,s),Z(s,t))ds-\int_t^T Z(t,s)dW(s)\\
&= \psi(t)+\int_{T-\delta}^T g(t,s,Y(s),Z(t,s),Z(s,t))ds-\int_{T-\delta}^T Z(t,s)dW(s)\\
&\quad +\int_t^{T-\delta} g(t,s,Y(s),Z(t,s),Z(s,t))ds-\int_t^{T-\delta} Z(t,s)dW(s)\\
&\equiv \widehat{\psi}(t)+\int_t^{T-\delta} g(t,s,Y(s),Z(t,s),Z(s,t))ds-\int_t^{T-\delta} Z(t,s)dW(s),
\end{aligned}
$$

where it is not obvious if $\widehat{\psi}(t)$ is/can be chosen $\mathcal{F}_{T-\delta}$-measurable!

4. Properties of Solutions.

- *A Duality Principle*

ODE case: Consider

$$\dot{x}(t) = Ax(t) + f(t), \quad x(0) = 0, \tag{4.1}$$

$$\dot{y}(t) = -A^T y(t) - g(t), \quad y(T) = 0. \tag{4.2}$$

Then

$$\frac{d}{dt}\big[\,\langle\, x(t), y(t)\,\rangle\,\big] = \langle\, f(t), y(t)\,\rangle - \langle\, x(t), g(t)\,\rangle\,.$$

Thus,

$$\int_0^T \langle\, x(t), g(t)\,\rangle\, dt = \int_0^T \langle\, y(t), f(t)\,\rangle\, dt. \tag{4.3}$$

- (4.2) is called an adjoint equation of (4.1).
- (4.3) is called a duality between (4.1) and (4.2).
- (linear) SDE and BSDE have a similar duality principle. Itô's formula is commonly used.

Theorem 4.1. Let $\varphi \in L^2_{\mathbb{F}}(0, T)$ and $\psi \in L^2((0, T) \times \Omega)$. Let

$$X(t) = \varphi(t) + \int_0^t A_0(t, s)X(s)ds + \int_0^t A_1(t, s)X(s)dW(s), \tag{4.4}$$

$$\begin{aligned} Y(t) = \psi(t) + \int_t^T \big[A_0(s, t)^T Y(s) + A_1(s, t)^T Z(s, t)\big] ds \\ - \int_t^T Z(t, s)dW(s), \quad t \in [0, T]. \end{aligned} \tag{4.5}$$

Then the following relation holds:

$$\mathbb{E} \int_0^T \langle Y(t), \varphi(t) \rangle \, dt = \mathbb{E} \int_0^T \langle \psi(t), X(t) \rangle \, dt. \tag{4.6}$$

(4.5) — the adjoint equation of (4.4)
(4.6) — the duality between (4.4) and (4.5).

- *A Comparison Theorem*

Consider BSDEs: $(k = 1, 2)$

$$(4.7)\qquad \begin{cases} dY^k(t) = -g^k(t, Y^k(t), Z^k(t))dt + Z^k(t)dW(t), \\ Y^k(T) = \xi^k. \end{cases}$$

Let

$$(4.8)\qquad \begin{cases} g^1(t, s, y, z) \le g^2(t, s, y, z), \quad \forall (t, s, y, z), \\ \xi^1 \le \xi^2, \quad \text{a.s.} \end{cases}$$

Then

$$(4.9)\qquad Y^1(t) \le Y^2(t), \qquad t \in [0, T], \text{ a.s.}$$

- Itô formula is used in the proof.
- Does not rely on the comparison of FSDEs.

Theorem 4.2. For $k = 1, 2$, let $g^k : [0, T]^2 \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ and $\psi^k(\cdot) \in L^2_{\mathbb{F}}(0, T; \mathbb{R})$ such that

$$\begin{cases} g^1(t, s, y, \zeta) \le g^2(t, s, y, \zeta), \quad \forall (t, s, y, \zeta), \\ \psi^1(t) \le \psi^2(t), \qquad t \in [0, T], \text{ a.s.} \end{cases} \tag{4.10}$$

Let $(Y^k(\cdot), Z^k(\cdot\,,\cdot))$ be the adapted M-solution of BSIVE

$$\begin{aligned} Y^k(t) = \psi^k(t) + \int_t^T g^k(t, s, Y^k(s), Z^k(s, t))ds \\ - \int_t^T Z^k(t, s)dW(s). \end{aligned} \tag{4.11}$$

Then the following holds:

$$Y^1(t) \le Y^2(t), \qquad \forall t \in [0, T]. \tag{4.12}$$

• *Sub-Additivity and Convexity.*

Let $(Y(\cdot), Z(\cdot\,,\cdot))$ be the adapted solution of BSVIE

$$
\begin{aligned}
Y(t) = \psi(t) &+ \int_t^T g(t,s,Y(s),Z(s,t))ds \\
&- \int_t^T Z(t,s)dW(s).
\end{aligned}
\tag{4.13}
$$

Denote

$$
\rho(t;\psi(\cdot)) = Y(t), \qquad t \in [0,T]. \tag{4.14}
$$

• $\psi(\cdot) \mapsto \rho(t;-\psi(\cdot))$ is essentially a *dynamic risk measure*.

Proposition 4.4. Let $g : [0, T]^2 \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$.
(i) Suppose $(y, \zeta) \mapsto g(t, s, y, \zeta)$ is sub-additive:

$$g(t, s, y_1 + y_2, \zeta_1 + \zeta_2) \le g(t, s, y_1, \zeta_1) + g(t, s, y_2, \zeta_2),$$
$$\forall (t, s) \in [0, T]^2,\ y_1, y_2 \in \mathbb{R},\ \zeta_1, \zeta_2 \in \mathbb{R}^d,\ \text{a.s.}\ ,$$

Then $\psi(\cdot) \mapsto \rho(t; \psi(\cdot))$ is sub-additive:

$$\rho(t; \psi_1(\cdot) + \psi_2(\cdot)) \le \rho(t; \psi_1(\cdot)) + \rho(t; \psi_2(\cdot)), \quad t \in [0, T],\ \text{a.s.}$$

(ii) Suppose $(y,z)\mapsto g(t,s,y,\zeta)$ is convex:

$$\begin{aligned}&g(t,s,\lambda y_1+(1-\lambda)y_2,\lambda\zeta_1+(1-\lambda)\zeta_2)\\&\le\lambda g(t,s,y_1,\zeta_1)+(1-\lambda)g(t,s,y_2,\zeta_2),\\&\qquad\forall(t,s)\in[0,T]^2,\ y_1,y_2\in\mathbb{R},\ \zeta_1,\zeta_2\in\mathbb{R}^d,\ \text{a.s.}\ ,\quad\lambda\in[0,1].\end{aligned}$$

Then $\psi(\cdot)\mapsto\rho(t;\psi(\cdot))$ is convex:

$$\rho(t;\lambda\psi_1(\cdot)+(1-\lambda)\psi_2(\cdot))\le\lambda\rho(t;\psi_1(\cdot))+(1-\lambda)\rho(t;\psi_2(\cdot)),$$
$$t\in[0,T],\ \text{a.s.}\ ,\ \lambda\in[0,1].$$

• Similar results hold if exchanging super-additivity and sub-additivity, convexity and concavity, respectively.

5. Some Remarks:

• Regularity of adapted M-solutions:

$$(1.6)\qquad \begin{aligned} Y(t) = \psi(t) &+ \int_t^T g(t,s,Y(s),Z(t,s),Z(s,t))ds \\ &- \int_t^T Z(t,s)dW(s), \qquad t \in [0,T]. \end{aligned}$$

Continuity of $t \mapsto Y(t)$ is not trivial. Malliavin calculus will be involved.

• Necessary conditions for optimal control of FSVIEs can be obtained.

• Existence of dynamic risk measure for general position processes.

……

Thank You!