BSDEs and their connection to stochastic control, Convex BSDEs

Joscha Diehl

December 17, 2008

1 Introduction

This talk will cover two topics.

Firstly I will show that a certain stochastic control problem is connecte to a LBSDE. From that one can derive properties of the solution to the control problem.

Secondly I will show, that Convex BSDEs, i.e. BSDEs with convex standard generator f, can be computed as the supremum of certain linear BSDEs, for which we already know how to compute the solution.

2 Review

In the previous talk the following results was derived.

Theorem 2.1 (Linear BSDE)**.** *Let (β, γ) be a bounded $(\mathbb{R}, \mathbb{R}^n)$-valued predictable process, $\phi \in \mathbb{H}^2_T(\mathbb{R})$ and $\xi \in \mathbb{L}^2_T(\mathbb{R})$. Then the LBSDE*

$$-dY_t = [\phi_t + Y_t\beta_t + Z_t^*\gamma_t]dt - Z_t^* dW_t, \qquad Y_T = \xi$$

has a unique solution (Y, Z) in $\mathbb{H}^2_{T,\beta}(\mathbb{R}) \times \mathbb{H}^2_{T,\beta}(\mathbb{R}^n)$ and Y_t is given by

$$Y_t = E[\xi\Gamma_T^t + \int_t^T \Gamma_s^t \phi_s ds | \underset{t}{\mathcal{F}}] \quad \mathbb{P}\, a.s.$$

where Γ_s^t is the ***adjoint process*** *defined for $s \geq t$ by the (forward) LSDE*

$$d\Gamma_s^t = \Gamma_s^t[\beta_s ds + \gamma_s^* dW_s], \qquad \Gamma_t^t = 1$$

Theorem 2.2 (Comparison Theorem)**.** *Let (f^1, ξ^1) and (f^2, ξ^2) be standard parameters of BSDEs with corresponding square-integrable solutions (Y^1, Z^1) and (Y^2, Z^2). Suppose that*

$$\xi^1 \geq \xi^2 \quad \mathbb{P}\, a.s.$$
$$f^1(t, Y_t^2, Z_t^2) - f^2(t, Y_t^2, Z_t^2) \geq 0 \quad d\,\mathbb{P} \times dt\ a.s.$$

Then we have almost surely

$$Y_t^1 \geq Y_t^2 \quad \forall t \in [0, T]$$

3 Stochastic Control

Consider the following **stochastic control** problem.

Let $\mathcal{U} := \{u : \Omega \times [0, T] \to U :$ u is predictable $\}$ be the set of feasible **controls**, where U is a Polish space. Let $\{H^u\}_{u \in \mathcal{U}}$ be a family of processes, defined via the the SDEs

$$dH_t^u = H_t^u[d(t, u_t)dt + n(t, u_t)dW_t], \qquad H_0^u = 1$$

where $d(t,u)$ and $n(t,u)$ are predictable processes (for fixed u) uniformly bounded by δ_t and μ_t respectively.

The problem is to minimize, over all feasible controls, the **objective function**

$$J(u) := E[\int_0^T H_t^u k(t,u_t)dt + H_T^u K(u_T)]$$

where

- $K(\cdot,u_T)$ is the terminal condition, where $K(w,u)$ is assumed to be measurable with repect to $\mathcal{F}_T \times \mathcal{B}(\mathcal{U})$ and bounded by a square-integrable variable χ
- $k(\cdot,t,u_t)$ is the running cost associated with the control process u, where $k(w,t,u)$ is assumed to be measurable with respect to $\mathbb{P} \times \mathcal{B}(\mathcal{U})$ and bounded by a square-integrable process $(\kappa_t : t \in [0,T])$
- δ, μ, κ and χ are supposed to be bounded

I will now show, that the solution of this problem is connected to solving a LBSDE.

With

$$f^u(t,y,z) = k(t,u_t) + d(t,u_t)y + n(t,u_t)^* z \qquad \xi^u = K(u_T)$$

we can see that for the process (Y^u, Z^u) that solves the LBSDE (f^u, ξ^u):

$$J(u) = Y_0^u$$

by Theorem (2.1). H^u corresponds to the adjoint process associated with (Y^u, Z^u).

Now, to minimize the objective function, the follwing theorem is central.

Theorem. *(Verification Theorem) The parameters (f,ξ) defined by*

$$f(t,y,z) := \operatorname{ess\,inf}\{f^u(t,y,z) : u \in \mathcal{U}\}, \qquad \xi := \operatorname{ess\,inf}\{\xi^u : u \in \mathcal{U}\}$$

are standard parameters. Let (Y,Z) be the solution of the BSDE associated with terminal condition ξ. Then:

> *Y is the value function Y^* of the control problem; that is for each $t \in [0,T]$*
> $$Y_t = Y_t^* = \operatorname{ess\,inf}\{Y_t^u : u \in \mathcal{U}\}$$

For the proof, we are going to need the following

Lemma. *Let (f,ξ) and (f^α, ξ^α) be a family of standard parameters and let (X,Y) and (X^α, Y^α) be the solution of associated BSDEs. Suppose that*

- *the generators f^α are equi-Lipschitz with the same constant C, i.e. $d\,\mathbb{P} \times dt$ a.s.:*

$$\forall \alpha : |f^\alpha(w,t,y_1,z_1) - f^\alpha(w,t,y_2,z_2)| \le C(|y_1 - y_2| + |z_1 - z_2|)$$

- *for each $\epsilon > 0$, there exists a control α^ϵ s.t.*

$$f(t, Y_t, Z_t) = \operatorname{ess\,inf} f^\alpha(t, Y_t, Z_t) \geq f^{\alpha^\epsilon}(t, Y_t, Z_t) - \epsilon \quad d\,\mathbb{P} \times dt \ a.s. \qquad (1)$$
$$\xi = \operatorname{ess\,inf} \xi^\alpha \geq \xi^{\alpha^\epsilon} - \epsilon \quad \mathbb{P} \ a.s.$$

Then

$$Y_t = \operatorname{ess\,inf} Y_t^\alpha \forall t \quad \in [0, T], \qquad \mathbb{P} \ a.s.$$

Proof Of The Lemma. Let $\epsilon > 0$ be given, fix an α^ϵ as given by the assumptions. Let $d = n = 1$, i.e. $Y : \Omega \times [0, T] \to \mathbb{R}$ and $Z : \Omega \times [0, T] \to \mathbb{R}$. Define

$$\delta Y_t := Y_t - Y_t^{\alpha^\epsilon}$$
$$\delta Z_t := Z_t - Z_t^{\alpha^\epsilon}$$

Then

$$\begin{aligned}
-d\delta Y_t &= -dY_t - (-dY_t^{\alpha^\epsilon}) \\
&= [f(t, Y_t, Z_t) - f^{\alpha^\epsilon}(t, Y_t^{\alpha^\epsilon}, Z_t^{\alpha^\epsilon})]dt - [Z_t - Z_t^{\alpha^\epsilon}]dW_t \\
&= [f(t, Y_t, Z_t) - f^{\alpha^\epsilon}(t, Y_t, Z_t) + f^{\alpha^\epsilon}(t, Y_t, Z_t) - f^{\alpha^\epsilon}(t, Y_t, Z_t^{\alpha^\epsilon}) + f^{\alpha^\epsilon}(t, Y_t, Z_t^{\alpha^\epsilon}) - f^{\alpha^\epsilon}(t, Y_t^{\alpha^\epsilon}, Z_t^{\alpha^\epsilon})]dt \\
&\quad - [Z_t - Z_t^{\alpha^\epsilon}]dW_t \\
&= [\phi_t + \beta_t \delta Y_t + \gamma_t \delta Z_t]dt - \delta Z_t dW_t \\
\delta Y_T &= \xi - \xi^{\alpha^\epsilon}
\end{aligned}$$

where

$$\phi_t := f(t, Y_t, Z_t) - f^{\alpha^\epsilon}(t, Y_t, Z_t)$$
$$\beta_t := \frac{f^{\alpha^\epsilon}(t, Y_t, Z_t^{\alpha^\epsilon}) - f^{\alpha^\epsilon}(t, Y_t^{\alpha^\epsilon}, Z_t^{\alpha^\epsilon})}{\delta Y_t}$$
$$\gamma_t := \frac{f^{\alpha^\epsilon}(t, Y_t, Z_t) - f^{\alpha^\epsilon}(t, Y_t, Z_t^{\alpha^\epsilon})}{\delta Z_t}$$

Since f is Lipschitz, β, γ are bounded. Also, ϕ is in $\mathbb{H}_T^2$. Therefore we have a LBSDE as in Theorem (2.1). We derive, that

$$\delta Y_t = E[\int_t^T \Gamma_{t,s} \phi_s ds + \Gamma_{t,T} \delta Y_T | \underset{t}{\mathcal{F}}]$$

where $\Gamma_{t,\cdot}$ is the adjoint process of the above LBSDE, i.e. it solves the following SDE

$$d\Gamma_{t,s} = \Gamma_{t,s}[\beta_s ds + \gamma_t dWs], \quad \Gamma_{t,t} = 1$$

Hence, we have

$$0 \geq \delta Y_t \overset{(*)}{\geq} -\epsilon E[\int_t^T \Gamma_{t,s} ds + \Gamma_{t,T} | \underset{t}{\mathcal{F}}] \overset{(**)}{\geq} -\epsilon(T + 1)e^{CT}$$

Where $(*)$ follows from (1).

For $(**)$ we consider only the case $t = 0$, the general case is similar. Firstly, the SDE for $\Gamma_s := \Gamma_{0,s}$ has an explicit solution:

$$\Gamma_s = \exp[\int_0^s (\beta_v - \frac{1}{2}\gamma_v)dv + \int_0^s \gamma_v dW_v]$$

Hence we see, that Γ_s is (a.s.) positive. Hence $\forall s \in [0, T]$

$$\begin{aligned} 0 \leq E[\Gamma_s] &= 1 + E[\int_0^s \Gamma_v \beta_v dv] + E[\int_0^s \Gamma_v \gamma_v dW_v] \\ &= 1 + \int_0^s E[\Gamma_v \beta_v] dv \leq 1 + \int_0^s E[\Gamma_v] C dv \end{aligned}$$

Where C is the Lipschitz constant of the family f^{α^ϵ}.

Also, $E[\Gamma_s]$ is continuous in s. Indeed, because Γ_s has a.s. continuous paths [see [1] Theorem 5.2.9]

$$\Gamma_s \underset{s \to t}{\longrightarrow} \Gamma_t \quad a.s.$$

Now, by Burkholder-Davis-Gundy inequality [see [1] Theorem 3.3.28]

$$E[(\sup_{0 \leq v \leq t} \Gamma_v)^2] \leq K E[\langle \Gamma \rangle_t]$$

and

$$\begin{aligned} E[\langle \Gamma \rangle]_t &= E[\int_0^t \Gamma_v^2 \beta_v^2 dv] \\ &= \int_0^t E[\Gamma_v^2 \beta_v^2] dv \qquad \leq \int_0^t E[\Gamma_v^2 C^2] dv \leq \int_0^t D e^{Dv} C^2 dv < \infty \end{aligned}$$

Where the last inequality follows for some $D > 0$ from the bound on the L^2-norm of solutions to SDEs with Lipschitz coefficients [see [1] Theorem 5.2.9]. So $\sup_{0 \leq v \leq t} \Gamma_v$ is square-integrable (hence integrable) and majorizes all Γ_s. Hence by dominated convergence:

$$E[\Gamma_s] \underset{s \to t}{\longrightarrow} E[\Gamma_t]$$

Hence, Gronwalls Lemma yields

$$0 \leq E[\Gamma_s] \leq 1 + C \int_0^s \exp(C(s - v)) dv = \exp(Cs)$$

And therefore, $(**)$ is ok. □

Remark. *As pointed out to me by Professor Becherer, there is a much easier way to proof the last inequality:*

$$\begin{aligned} \Gamma_t &= \exp(\int_0^t [\beta_s - \frac{1}{2}\gamma_s] ds + \int_0^t \gamma_s dW s) \\ &\leq \exp(Ct) \exp(\int_0^t \gamma_s dW_s - \frac{1}{2}\gamma_s ds) \end{aligned}$$

The second multiplicand is a martingale, hence we get

$$E[\Gamma_t] \leq \exp(Ct)$$

We can now derive the Verification Theorem.

Proof of the Verification Theorem. Without proof (which only deals with measurability concerns), I assume that:

- f as defined in the statement is indeed a standard parameter

Now

- for every $\epsilon > 0$ there exists a feasible control u^ϵ s.t.

$$f(t, Y_t, Z_t) = \operatorname{ess\,inf} f^u(t, Y_t, Z_t) \geq f^{u^\epsilon}(t, Y_t, Z_t) - \epsilon, \quad d\,\mathbb{P} \times dt \quad a.s.$$
$$\xi = \operatorname{ess\,inf} \xi^u \geq \xi^{u^\epsilon} - \epsilon \quad a.s.$$

Let $\varepsilon > 0$ be given. For $(w, t) \in \Omega \times [0, T)$ fixed, the sets

$$\{u \in \mathbb{R} : f(t, Y_t(w), Z_t(w)) \geq k(t, w, u) + d(t, w, u)Y_t(w) + n(t, w, u)Z_t(w) - \varepsilon\}$$
$$\{u \in \mathbb{R} : \xi(w) \geq K(w, u) - \varepsilon\}$$

are nonempty. Since Y and Z are predictable, a measurable selection theorem [see ([2]) Theorem III.44/45] yields that there exists a $\mathbb{R}$-valued predictable process u^ε s.t.

$$f(t, Y_t, Z_t) = \operatorname{ess\,inf} f^u(t, Y_t, Z_t) \geq f^{u^\varepsilon}(t, Y_t, Z_t) - \varepsilon \; d\,\mathbb{P} \otimes dt \; a.s.$$
$$\xi = \operatorname{ess\,inf} \xi^u \geq \xi^{u^\varepsilon} - \varepsilon$$

Now, the statement follows from the preceding Lemma. □

4 Convex BSDE as Supremum

Let $f(t, y, z)$ be a standard generator of a BSDE convex w.r.t. (y, z) and let $F(t, \beta, \gamma)$ be the **polar process** associated with f:

$$F(w, t, \beta, \gamma) := \inf_{(y,z) \in \mathbb{R} \times \mathbb{R}} [f(w, t, y, z) - \beta y - \gamma z]$$

Its **effective domain** is

$$\mathcal{D}_F^t := \{(w, t, \beta, \gamma \in \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R} : F(w, t, \beta, \gamma) > -\infty\}$$

And we have

$$f(w, t, y, z) = \sup\{F(w, t, \beta, \gamma) + \beta y + \gamma z : (\beta, \gamma) \in \mathcal{D}_F^t\}$$

Now let

$$f^{\beta,\gamma}(t, y, z) := F(t, \beta_t, \gamma_t) + \beta_t y + \gamma_t z$$

where (β, γ) are predictable processes.

For these to be standard generators, we can only allow processes belonging to

$$\mathcal{A} := \{(\beta, \gamma) \in \mathcal{P} : E \int_0^T F(t, \beta_t, \gamma_t)^2 dt < +\infty\}$$

Lemma. *There exists an* $(\bar{\beta}, \bar{\gamma}) \in \mathcal{A}$ *s.t.*

$$f(t, Y_t, Z_t) = f^{\bar{\beta},\bar{\gamma}}(t, Y_t, Z_t), \quad d\,\mathbb{P} \times dt \ a.s.$$

Proof. f is uniformly Lipschitz and concave, hence the supremum in the conjugacy relation is actually achieved (for fixed (w, t)).

$f(\cdot, Y, Z)$, Y and Z are predictable processes, hence by a measurable selection theorem [see ([2]) Theorem III.44/45] it follows, that there exists a pair of predictable (bounded) processes $(\bar{\beta}, \bar{\gamma})$ s.t.

$$f(t, Y_t, Z_t) = f^{\bar{\beta},\bar{\gamma}}(t, Y_t, Z_t), \quad d\,\mathbb{P} \times dt \ a.s.$$

By assumption $f(\cdot, Y_t, Z_t)$, Z and Y are square integrable and $\bar{\beta}$, $\bar{\gamma}$ are bounded. Hence $F(\cdot, \bar{\beta}, \bar{\gamma})$ belongs to $\mathbb{H}_T^{2,1}$.

Hence, the pair $(\bar{\beta}, \bar{\gamma})$ which achieves the infimum in the conjugacy relation, belongs to $\mathcal{A}$. □

It follows

Theorem. *Let* f *be a convex standard generator and* $f^{\bar{\beta},\bar{\gamma}}$ *be the associated linear standard generators, with*

$$f = \operatorname{ess\,sup}\{f^{\bar{\beta},\bar{\gamma}} : (\beta, \gamma) \in \mathcal{A}\}, \quad d\,\mathbb{P} \times dt \ a.s.$$

Then $\mathbb{P}$*-a.s.*

$$Y_t = \operatorname{ess\,sup}\{Y_t^{\beta,\gamma} : (\beta, \gamma) \in \mathcal{A}\}$$

Proof.
- Because of the Comparison Theorem we have for every (β, γ): $Y_t \geq Y_t^{\beta,\gamma}$. Hence $Y_t \geq \operatorname{ess\,sup}\{Y_t^{\beta,\gamma}\}$.
- From the uniqueness theorem of BSDEs we have that $Y_t = Y_t^{\bar{\beta},\bar{\gamma}}$. Hence $\operatorname{ess\,sup}\{Y_t^{\beta,\gamma}\} \geq Y_t^{\bar{\beta},\bar{\gamma}} = Y_t \geq \operatorname{ess\,sup}\{Y_t^{\beta,\gamma}\}$.

□

4.1 Example

Consider the problem of replicating a contingent claim, when the rate R_t for borrowing is higher then the interest rate r_t. This leads to the following BSDE with convex generator:

$$\begin{aligned} dY_t &= r_t Y_t dt + \pi_t \sigma_t \theta_t dt + \pi_t \sigma_t dW_t - (R_t - r_t)(Y_t - \pi_t)^- dt \\ Y_T &= \xi \end{aligned} \tag{2}$$

Like all other coefficients, assume R_t to be bounded.

The generator is given by

$$f(t, y, z) = -r_t y - z\theta_t + (R_t - r_t)(y - \sigma_t^{-1} z)^-$$

Which leads to the associated polar process

$$F(t,\beta,\gamma) = \begin{cases} 0 & \text{if } \gamma = -\theta_t - \sigma_t^{-1}(r_t+\beta) \text{ and } r_t \le -\beta \le R_t \\ -\infty & \text{else} \end{cases}$$

Let us look at the first case, the other follow similarly. Then

$$\begin{aligned}
&\inf_{y,z}[-r_t y - z\theta_t + (R_t - r_t)(y - \frac{1}{\sigma_t}z)^- - \beta y - \gamma z] \\
&= \inf_{y,z}[-r_t y - z\theta_t + (R_t - r_t)(y - \frac{1}{\sigma_t}z)^- - \beta y + \theta_t z + \frac{1}{\sigma_t}(r_t+\beta)z] \\
&= \inf_{y,z}[(R_t - r_t)(y - \frac{1}{\sigma_t}z)^- + (-\beta - r_t)y + \frac{1}{\sigma_t}(r_t+\beta)z]
\end{aligned}$$

For $y \ge \frac{z}{\sigma_t}$ we get

$$\begin{aligned}
&\inf_{y,z}[(R_t - r_t)(y - \frac{1}{\sigma_t}z)^- + (-\beta - r_t)y + \frac{1}{\sigma_t}(r_t+\beta)z] \\
&= \inf_{y,z}[(-\beta - r_t)y + \frac{1}{\sigma_t}(r_t+\beta)z] \\
&\ge \inf_{y,z}[(-\beta - r_t)\frac{z}{\sigma_t} + \frac{1}{\sigma_t}(r_t+\beta)z] = 0
\end{aligned}$$

And for $y < \frac{z}{\sigma_t}$ we get

$$\begin{aligned}
&\inf_{y,z}[(R_t - r_t)(y - \frac{1}{\sigma_t}z)^- + (-\beta - r_t)y + \frac{1}{\sigma_t}(r_t+\beta)z] \\
&= \inf_{y,z}[(R_t - r_t)(\frac{1}{\sigma_t}z - y) + (-\beta - r_t)y + \frac{1}{\sigma_t}(r_t+\beta)z] \\
&= \inf_{y,z}[\frac{z}{\sigma_t}(R_t+\beta) + y(-R_t - \beta)] \\
&\ge \inf_{y,z}[y(R_t+\beta) + y(-R_t - \beta)] = 0
\end{aligned}$$

So with this polar process we have by the Theorem, that the unique solution $(Y, \sigma\pi)$ of the BSDE (2) satisfies

$$Y_t = \operatorname{ess\,sup}\{Y_t^\beta : r_t \le \beta_t \le R_t\}$$

where

$$\begin{aligned}
-dY_t^\beta &= [-\beta_t Y_t^\beta - \sigma_t\theta_t + (r_t - \beta_t)\pi_t^\beta]dt - \pi_t^\beta \sigma_t dW_t \\
Y_T &= \xi
\end{aligned}$$

Appendix A Background

A.1 Definitions / Theorems of the El Karoui paper

Definition. (f, ξ) are said to be *standard parameters* for the BSDE if

- $\xi : \Omega \to \mathbb{R}^d$ is F_T-measurable and in $L^2_T(\mathbb{R}^d)$
- $f : \Omega \times \mathbb{R}^x \times \mathbb{R}^d \times \mathbb{R}^{n\times d}$ is $P \times \mathbb{B}^d \times \mathbb{B}^{n\times d}$-measurable, $f(\cdot, 0, 0) \in \mathbb{H}^2_T(\mathbb{R}^d)$ and f is uniformly Lipschitz; i.e. there exists $C > 0$ s.t. $dP \times dt$ a.s.

$$|f(w,t,y_1,z_1)-f(w,t,y_2,z_2)| \leq C(|y_1-y_2|+|z_1-z_2|) \qquad \forall (y_1,z_1),(y_2,z_2)$$

Definition. Let X be a random variable. Then the **essential infimum** of X is defined as

$$\operatorname{ess\,inf} X := \sup\{a \in \mathbb{R} : P(X < a) = 0\}$$

Let X_t, Y_t be random processes. Then X_t **minorizes** Y_t if

$$P(\omega : \exists t \in [0,T] \text{ s.t. } X_t > Y_t) = 0$$

We say

$$X_t = \operatorname*{ess\,inf}_\alpha Y^\alpha_t$$

if for every α, X_t minorizes Y^α_t and every process which does this also, minorizes X_t. (existence for RCLL processes: Dellacherie)

For $T = 0$ this gives the ess inf of a family of random variables.

A.2 Other results used

Theorem (Gronwall's Lemma)**.** *Let $g : \mathbb{R}_+ \to \mathbb{R}$ be a continuous function s.t. for some constants α and β*

$$0 \leq g(t) \leq \alpha + \beta \int_0^t g(s)ds \quad \forall t \geq 0$$

Then

$$g(t) \leq \alpha e^{\beta t}$$

B Bibliography

References

[1] Ioannis Karatzas, Steven E. Shreve
Brownian Motion and Stochastic Calculus, 2nd Edition
Springer-Verlag, 1988

[2] Claude Dellacherie, Paul-Andre Meyer
Probalites et potentiel, Chapitres I a IV
Hermann, 1975