Risk measures and BSDEs

H. Mete Soner,
ETH Zürich and Swiss Finance Institute

Stochastic Models and Control
March 29, 2011

Collaborators

This is part of joint program with

- Nizar Touzi, Ecole Polytechnique, Paris,
- Jianfeng Zhang, USC,

we have collaborated with

- Umut Çetin, London School of Economics,
- Patrick Cheridito, Princeton,
- Yan Dolinsky, ETH,
- Selim Gökay, ETH,
- Marcel Nutz, ETH.

Outline

Outline

Mathematical Framework

To express this problem mathematically, as usual, we assume that we are given a finite maturity $T > 0$, a probability space $(\Omega, \mathbb{P}, \mathbb{F})$ where $\mathbb{P}$ is a probability measure and the filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ is an increasing sequence of σ-algebras.

Then the risky financial position is simply a $\mathcal{F}_T$ measurable random variable $\xi \in L^0$.

A risk measure is a map

$$\rho : D \subset L^0 \to \mathbb{R}^1.$$

Of course we need to place restrictions on this map.

Examples

- A simple (and not useful) example with $D = L^1(\Omega, \mathbb{P}, \mathcal{F}_T)$ is

$$\rho(\xi) = \mathbb{E}^{\mathbb{P}}[-\xi].$$

- More useful example is

$$\rho(\xi) = \sup_{\mathbb{Q} \in \mathcal{Q}} \left\{ \mathbb{E}^{\mathbb{Q}}[-\xi] \right\},$$

  where $\mathcal{Q}$ is a given set of probability measures. The choice of D is subtle. If all elements in $\mathcal{Q}$ are absolutely continuous with respect to $\mathbb{P}$, then we can use $D = L^\infty(\Omega, \mathbb{P}, \mathcal{F}_T)$.

Coherent/Convex Risk Measures

Artzner *et. al.* postulates that ρ should satisfy

1. (cash-invariance) $\rho(\xi + c) = \rho(\xi) - c, \quad \forall c \in \mathbb{R}^1.$
2. (monotonicity) $\rho(\xi) \geq \rho(\hat{\xi}), \quad \text{if} \quad \xi \leq \hat{\xi}.$
3. (diversification) $\rho(\xi + \hat{\xi}) \leq \rho(\xi) + \rho(\hat{\xi}).$
4. (positive homogeneity) $\rho(\lambda\xi) = \lambda\rho(\xi), \quad \forall \lambda \geq 0.$

Any map satisfying above is a coherent risk measure.

Föllmer & Schied replaced the last condition by convexity to define the convex risk measures.

Continuity and Structure

Suppose that a convex risk measure satisfy an additional Fatou type continuity assumption :

$$\xi_n \to \xi \quad \text{in } D = L^\infty(\mathbb{P}) \quad \Rightarrow \quad \liminf \rho(\xi_n) \geq \rho(\xi).$$

Then, there is a convex subset $\mathcal{Q}$ of equivalent probability measures to $\mathbb{P}$ so that

$$\rho(\xi) = \sup_{\mathbb{Q} \in \mathcal{Q}} \left\{ \mathbb{E}^{\mathbb{Q}}[-\xi] + F(d\mathbb{Q}/d\mathbb{P}) \right\},$$

where F is a convex function.

Definition

For each $t \in [0, T]$, we consider a convex risk measure

$$\rho_t : D \subset L^0(\mathcal{F}_T) \to L^0(\mathcal{F}_t).$$

We only relax the cash-invariance by

$$\rho_t(\xi + \eta) = \rho_t(\xi) - \eta, \quad \forall \eta \in L^0(\mathcal{F}_t).$$

We also require time consistency :

$$\rho_t(\xi) = \rho_t(\rho_s(\xi)), \quad \forall t \leq s \leq T.$$

This condition is like the dynamic programming principle or like the semi-group property. There are technical issues and a viscosity type definition is better.

PDEs

Let W be the standard d-dimensional Brownian motion and the filtration be generated by it. Assume a Markovian structure that $\xi = \varphi(x + W_T)$, where φ is a bounded, Lipschitz, smooth, deterministic function. Let u be the bounded, Lipschitz, smooth, solution of

$$-u_t(t,x) - \frac{1}{2}\Delta u(t,x) + H(t, Du(t,x)) = 0, \quad (t,x) \in (0,T) \times \mathbb{R}^d,$$

with terminal condition

$$u(T,x) = -\varphi(x), \quad x \in \mathbb{R}^d.$$

Here the nonlinearity H is a given deterministic function, convex in the gradient variable.

PDEs continued

Then, one can check that $\rho_t(\xi) := u(t, x + W_t)$ is a dynamic risk measure. Indeed,

$$H(t,p) = \sup\{-\alpha \cdot p + L(t,\alpha)\}$$

$$\begin{aligned} u(t,x) &= -\sup_{\alpha.} \mathbb{E}^{\mathbb{P}} \left[\varphi \left(x + W_T - W_t + \int_t^T \alpha_s ds \right) + \int_t^T L(s,\alpha_s) ds \right] \\ &= -\sup_{\alpha.} \left\{ \mathbb{E}^{\mathbb{Q}^\alpha} \left[\varphi(x + W_T^\alpha - W_t^\alpha) + \int_t^T \frac{d\mathbb{P}}{d\mathbb{Q}^\alpha} L(s,\alpha_s) ds \right] \right\}, \end{aligned}$$

where $\mathbb{Q}^\alpha$ is the measure under which

$$W^\alpha := W_t + \int_t^T \alpha_s ds$$

is a Brownian motion (i.e., Girsanov).

Conditional Expectations

In the coherent case, we want to extend the static risk measure

$$\rho(\xi) = \rho_0(\xi) = \sup_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}^{\mathbb{Q}}[-\xi].$$

In the PDE situation, $\rho_t(\xi) = u(t, x + W_t)$ does the job. For a non-Markov ξ, the natural definition is,

$$\rho_t(\xi) = \operatorname*{esssup}_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}^{\mathbb{Q}}\left[-\xi \mid \mathcal{F}_t\right].$$

Important : To define esssup, we need to fix the null sets or equivalently fix a probability measure.
In this case, the elements in $\mathcal{Q}$ are all equivalent to each other and in particular to $\mathbb{P}$. So one can easily choose $\mathbb{P}$ to define the essential supremum.

Outline

Fully nonlinear PDE

Again for a Markov $\xi = \varphi(x + W_T)$, we can use the scalar PDE

$$-u_t + H(t, x, Du(t,x), D^2u(t,x)) = 0,$$

with data $u(T,x) = -\varphi(x)$. Here H is a given parabolic, convex nonlinearity. Then, as before, $\rho_t(\xi) = u(t, x + W_t)$ is a convex, time consistent risk measure but does not have the Fatou property..

Fully nonlinear PDE

Again for a Markov $\xi = \varphi(x + W_T)$, we can use the scalar PDE

$$-u_t + H(t, x, Du(t, x), D^2 u(t, x)) = 0,$$

with data $u(T, x) = -\varphi(x)$. Here H is a given parabolic, convex nonlinearity. Then, as before, $\rho_t(\xi) = u(t, x + W_t)$ is a convex, time consistent risk measure but does not have the Fatou property..

To go beyond the Markov structure :

In joint work with Cheredito, S., Touzi, Victoir (CPAM, 2002), we also proved a stochastic representation for this PDE. This can be used to extend the definition. This is done in joint work with Touzi, Zhang (SPA-2011, EJP-2011, PTRF-2011).

Special nonlinearity

S. Peng takes the special nonlinearity

$$H(t, x, Du, D^2u) = G(D^2u) := \sup_{\underline{a} \leq a \leq \overline{a}} \left[-a : D^2u \right],$$

where $0 \leq \underline{a} \leq \overline{a}$ are given bounds. In financial terms, this means that the volatility of the stock is only known to be in $[\sqrt{2\underline{a}}, \sqrt{2\overline{a}}]$. This is the uncertain volatility model and the Markov case was studied by Avallaneda, Levy and Paras and by Lyons in 1995. It is clear that the solution is the value function of an optimal control problem in which the state process is

$$dX_t = \sigma_t dW_t,$$

where we can choose σ_t in the interval $[\sqrt{2\underline{a}}, \sqrt{2\overline{a}}]$.

Non-markov case

Intuitively, the value function interpretation implies that

$$\mathbb{E}^G(-\xi) = \mathbb{E}_0^G(-\xi) = \sup_{\mathbb{Q}\in\mathcal{P}} \mathbb{E}^{\mathbb{Q}}[-\xi],$$

where $\mathbb{Q} \in \mathcal{P}$ iff the canonical map has quadratic variation under $\mathbb{Q}$ is in the given interval. I will make this clear soon.

This definition does not require any Markov Structure.

Non-markov case

Intuitively, the value function interpretation implies that

$$\mathbb{E}^G(-\xi) = \mathbb{E}_0^G(-\xi) = \sup_{\mathbb{Q}\in\mathcal{P}} \mathbb{E}^{\mathbb{Q}}[-\xi],$$

where $\mathbb{Q} \in \mathcal{P}$ iff the canonical map has quadratic variation under $\mathbb{Q}$ is in the given interval. I will make this clear soon.

This definition does not require any Markov Structure.

What we want more is a time consistent extension, or equivalently, a nonlinear conditional expectation. Indeed, Peng sees $\mathbb{E}^G$ as nonlinear expectation and would like to define $\mathbb{E}_t^G$.

This was done in the semilinear case by suing the backward stochastic differential equations of Pardoux & Peng.

Warning

Since I am interested in the expected value,
from now on, I will define things for ξ not for $-\xi$.

Also I will take $T = 1$.

$\xi = \varphi(B_1)$

Assume that φ is bounded and Lipschitz continuous on $\mathbb{R}^d$. Let u be the unique bounded Lipschitz viscosity solution of the following parabolic equation,

$$-u_t - G(D^2 u) = 0 \quad \text{on } [0,1), \text{ and } u(1,x) = \varphi(x).$$

Then, Peng defines the conditional G-expectation of the random variable $\varphi(B_1)$ at t by

$$\mathbb{E}_t^G\left[\varphi(B_1)\right] := u\left(t, B_t\right).$$

In particular, the G-expectation of $\varphi(B_1)$ is given by

$$\mathbb{E}^G[\varphi(B_1)] := \mathbb{E}_0^G[\varphi(B_1)] = u(0,0).$$

$\xi = \varphi(B_{t_1}, \ldots, B_1)$

We essentially proceed the same way. Indeed, in the interval (t_{i-1}, t_i), let

$$\mathbb{E}_t^G[\xi] = \mathbb{E}_t^G[\varphi(B_{t_1}, \ldots, B_{t_n})] := v_i(t, B_{t_1}, \ldots, B_{t_{i-1}}, B_t),$$

where $\{v_i\}_{i=1,\ldots,n-1}$ is the unique, bounded, Lipschitz viscosity solution of the following equation,

$$\begin{aligned} -\partial_t v_i - G\left(D^2 v_i\right) &= 0, \qquad t_{i-1} \le t < t_i \qquad \text{and} \\ v_i\left(t_i, x_1, \ldots, x_{i-1}, x\right) &= v_{i+1}\left(t_i, x_1, \ldots, x_{i-1}, x, x\right), \end{aligned}$$

and v_n solves the above equation with final data $v_n(1, x_1, \ldots, x_{n-1}, x) = \varphi(x_1, \ldots, x_{n-1}, x)$.

The space $\mathcal{L}^p_G$

Let $\mathcal{L}_{ip}$ be the space of all random variables $\xi = \varphi(B_{t_1}, \ldots, B_1)$ with a bounded and Lipschitz φ. Then on $\mathcal{L}_{ip}$ we have just defined $\mathbb{E}^G_t$. Therefore, on $\mathcal{L}_{ip}$ we can also define a semi-norm

$$\|\xi\|^p_{\mathcal{L}^P_G} := \mathbb{E}^G[\ |\xi|^p\ ].$$

We then define the integrability class $\mathcal{L}^p_G$ as the closure of $\mathcal{L}_{ip}$ under this norm. Peng then defines G conditional expectations on $\mathcal{L}^p_G$ by a closure argument.

But these spaces do not even contain all bounded functions and require Lusin type of regularity (in fact, almost sure continuity).

More general nonlinearity

One may follow the Peng construction for a more general nonlinearity. Indeed, one may start with the nonlinear equation

$$-u_t + H(t, Du(t,x), D^2u(t,x)) = 0.$$

Then, follow the closure argument to obtain a similar nonlinear conditional expectations.

Set-up

Let $\Omega := C([0,1] : \mathbb{R}^d)$ and $B_t(\omega) = \omega_t$ be the canonical map.
In the classical theory, we consider only one measure which is the Wiener measure. However, in our model we consider all measures $\mathbb{Q}$ so that under this measure the canonical map is a martingale and the quadratic variation $\langle B \rangle_t$ satisfies

$$\underline{a}dt \leq d\langle B \rangle_t \leq \overline{a}dt,$$

where $0 \leq \underline{a} \leq \overline{a}$ are given volatility bounds.
We let $\mathcal{P}$ be the set of all such measures.
Note that, in general, measures in $\mathcal{P}$ are singular to each other.
We use the filtration generated by B without completing.

Problem of super-replication

Assume a bounded random variable $\xi \in \mathcal{F}_1$ is given. The minimal super-replication cost is given by

$$v(\xi) := \inf \left\{ x | \exists \; H \text{ s.t., } x + \int_0^1 H_s dB_s \geq \xi, \; \mathbb{P} - a.s., \; \forall \; \mathbb{Q} \in \mathcal{P} \right\}.$$

Problem of super-replication

Assume a bounded random variable $\xi \in \mathcal{F}_1$ is given. The minimal super-replication cost is given by

$$v(\xi) := \inf \left\{ x | \exists \ H \text{ s.t., } x + \int_0^1 H_s dB_s \geq \xi, \ \mathbb{P} - a.s., \ \forall \ \mathbb{Q} \in \mathcal{P} \right\}.$$

We also would like to construct the hedge H.

Problem of super-replication

Assume a bounded random variable $\xi \in \mathcal{F}_1$ is given. The minimal super-replication cost is given by

$$v(\xi) := \inf \left\{ x | \exists\ H \text{ s.t., } x + \int_0^1 H_s dB_s \geq \xi,\ \mathbb{P} - a.s.,\ \forall\ \mathbb{Q} \in \mathcal{P} \right\}.$$

We also would like to construct the hedge H.

Here the difficulty emanates from the fact that $\mathcal{P}$ contains uncountably many measures that are singular to each other.

Moreover, we want to describe the hedge for all measures and not only for the "optimal" one.

Description of the value

For any $\mathbb{Q} \in \mathcal{P}$ and $x > v(\xi)$, we have $x + \int_0^1 H_s dB_s \geq \xi$ for some Z. Take expected value $x \geq \mathbb{E}^{\mathbb{Q}}[\xi]$, forall $\mathbb{Q} \in \mathcal{P}$, and $x > v(\xi)$. Hence,

$$v(\xi) \geq \sup_{\mathcal{P}} \ \mathbb{E}^{\mathbb{Q}}[\xi].$$

Description of the value

For any $\mathbb{Q} \in \mathcal{P}$ and $x > v(\xi)$, we have $x + \int_0^1 H_s dB_s \geq \xi$ for some Z. Take expected value $x \geq \mathbb{E}^{\mathbb{Q}}[\xi]$, forall $\mathbb{Q} \in \mathcal{P}$, and $x > v(\xi)$. Hence,

$$v(\xi) \geq \sup_{\mathcal{P}} \ \mathbb{E}^{\mathbb{Q}}[\xi].$$

By analogy with the results of El Karoui & Quenez (1993) for super-replication in an incomplete market, it is proved by Denis & Martini and Denis, Hu & Peng (2009) that they are equal :

$$v(\xi) = \mathbb{E}^{G}[\xi] := \sup_{\mathbb{Q} \in \mathcal{P}} \ \mathbb{E}^{\mathbb{Q}}[\xi],$$

where $\mathbb{E}^{G}$ is the G-expectation of Peng.

Dual formula

For any $\xi \in \mathcal{L}^1_G$ and $t \in [0,1]$,

$$\mathbb{E}^G_t[\xi] = \operatorname*{esssup}_{\mathbb{P}' \in \mathcal{P}(t,\mathbb{Q})} \mathbb{E}^{\mathbb{P}'}[\,\xi \mid \mathcal{F}_t], \qquad \mathbb{P}-a.s., \ \forall\ \mathbb{Q} \in \mathcal{P},$$

$$\mathcal{P}(t,\mathbb{Q}) := \left\{\mathbb{P}' \in \mathcal{P} : \ \mathbb{P}' = \mathbb{Q} \text{ on } \mathcal{F}_t\right\}.$$

Dual formula

For any $\xi \in \mathcal{L}^1_G$ and $t \in [0,1]$,

$$\mathbb{E}^G_t[\xi] = \operatorname*{esssup}_{\mathbb{P}' \in \mathcal{P}(t,\mathbb{Q})} \mathbb{E}^{\mathbb{P}'}[\,\xi \mid \mathcal{F}_t]\,, \qquad \mathbb{P} - a.s., \ \forall\ \mathbb{Q} \in \mathcal{P},$$

$$\mathcal{P}(t,\mathbb{Q}) := \left\{ \mathbb{P}' \in \mathcal{P} : \ \mathbb{P}' = \mathbb{Q} \text{ on } \mathcal{F}_t \right\}.$$

Notice that tight hand side of the above depends on $\mathbb{Q}$. The technically difficult result is that the left hand side is independent of $\mathbb{Q}$. In other words the G-conditional expectation is an aggregation of the right hand side. Since these random variables are defined only upto $\mathbb{Q}$ null sets, the aggregation is hard.

Outline

Statement

Theorem

Given ξ, there are (Y_t, Z_t, K_t) satisfying for all $\mathbb{Q} \in \mathcal{P}$

$$\begin{aligned} Y_t := \mathbb{E}_t^G[\xi] &= \xi - \int_t^1 H_s dB_s + K_1 - K_t, \qquad \mathbb{Q} - a.s. \\ &= \mathbb{E}^G[\xi] + \int_0^t H_s dB_s - K_t, \qquad \mathbb{Q} - a.s., \end{aligned}$$

where $\mathbb{E}_t^G$ is the G-conditional expectation of Peng. Moreover, K_t is non-decreasing with $K_0 = 0$ and it is minimal in the sense that it is a G-martingale :

$$K_t = \mathbb{E}_t^G[K_s], \qquad 0 \leq t \leq s \leq 1.$$

References

- Theorem stated here is proved in a recent paper of ours, Soner, Touzi & Zhang (SPA-2011).
- Prior to our work a martingale representation was proved by Xu J. & Zhang B. (2009). They considered only symmetric G-martingales. N is a symmetric G-martingale if both N and $-N$ are G-martingales. And a symmetric G-martingale is a pure stochastic integral.
- Yongsheng Song Later, used our approach in all L^p spaces not only in L^2.

$\xi = \varphi(B_1)$

Assume that there is a smooth solution u of

$$-u_t - G(D^2 u) = 0 \text{ on } [0,1), \text{ and } u(1,x) = \varphi(x).$$

Then, we define on all of Ω

$$\begin{aligned} Y_t &:= \mathbb{E}_t^G[\xi] = u(t, B_t) \\ H_t &:= \nabla u(t, B_t) \\ K_t &:= \int_0^t \left(G(D^2 u(s, B_s)) - \frac{1}{2} \left[\hat{a}_s : D^2 u(s, B_s) \right] \right) ds, \end{aligned}$$

where $\hat{a}_t := d\langle B\rangle_t / dt$ is the universal quadratic variation process.

$\xi = \varphi(B_1)$ continued

By a direct application of the Ito's formula and the PDE, we conclude that for every $\mathbb{Q} \in \mathcal{P}$,

$$dY_t = H_t dB_t - dK_t, \qquad \mathbb{Q} - a.s..$$

Hence the martingale representation is proved once we show the properties of the process K. The definition of G yields

$$\begin{aligned} dK_t &= \left\{ G(D^2 u(s,B_s)) - \frac{1}{2}\left[\hat{a}_s : D^2 u(s,B_s)\right] \right\} dt \\ &= \frac{1}{2}\left[\sup_{\underline{a} \le a \le \overline{a}} (a : D^2 u(s,B_s)) - (\hat{a}_s : D^2 u(s,B_s)) \right] dt. \end{aligned}$$

Since for $\mathbb{Q} \in \mathcal{P}$, $\underline{a} \le \hat{a}_s \le \overline{a}$, $dK_t \ge 0$. For the optimal a, $dK_t = 0$.

In general

For $\xi = \varphi(B_{t_1}, \ldots, B_1)$, we use the same construction.

Then, we use the backward SDE theory to get a priori estimates and use them in a closure argument.

Estimates and the spaces need to be defined to perform this closure argument.

Open question

In the space $\mathcal{L}_{ip}$, the increasing process K has structure. Namely,

$$dK_t = \left\{ G(D^2 u(s, B_s)) - \frac{1}{2} \left[\hat{a}_s : D^2 u(s, B_s) \right] \right\} dt.$$

How can we generalize this to general random variables ?

Simpler question is : can we prove that K is absolutely continuous with respect to dt ?

Concluding...

1. Essentially, we studied the "BSDE" : find Y, H, Γ so that

$$\begin{aligned} dY_t &= G(\Gamma_t)dt - \frac{1}{2}\Gamma_t : d\langle B\rangle_t + H_t \cdot dB_t, \\ Y_T &= \xi, \qquad \mathbb{P} - a.s., \quad \forall \mathbb{P} \in \mathcal{P}. \end{aligned}$$

   This can be generalized to other nonlinearities.
2. The only important result from the PDE theory used is the uniqueness of viscosity solutions. But maybe this can be reversed to prove uniqueness by probabilistic techniques.
3. Jumps need to be included into the risk measures.
4. Numerical approaches using these ideas, in particular, Monte-Carlo techniques is important.

Thank you for your attention !