Dynamic programming approach to Principal-Agent problems

Thomas Bourany

Université Paris Dauphine

UPMC-Sorbonne Université

Problèmes variationnels et de transport en économie

Introduction

- This article recasts the Principal-Agent model with Moral Hazard in continuous time, using tools from stochastic calculus and stochastic control

Introduction

- This article recasts the Principal-Agent model with Moral Hazard in continuous time, using tools from stochastic calculus and stochastic control
- Follows a large literature in economics and mathematical finance :
  - Seminal contribution by Holmström and Milgrom
    - Agent effort influences the drift of a diffusion process
    - Happens 'as if' agent controlled the mean of a normal distribution
    - Optimal contract is linear in output

Introduction

- This article recasts the Principal-Agent model with Moral Hazard in continuous time, using tools from stochastic calculus and stochastic control
- Follows a large literature in economics and mathematical finance :
  - Seminal contribution by Holmström and Milgrom
    - Agent effort influences the drift of a diffusion process
    - Happens 'as if' agent controlled the mean of a normal distribution
    - Optimal contract is linear in output
  - Mathematical tools developed by (among others) : Cvitanić and Zhang (book) and other articles by D. Possamai and N. Touzi.
    - More advanced tools from stochastic calculus
    - Dynamic Programming, BSDE, Stochastic Max. Principle (FBSDE)

Introduction

- This article provides a systematic method to solve any problem of this kind :
  - Principal observes fluctuations in output and offers a compensation scheme at terminal time.
  - Agent control the drift *and* the volatility of this output
  - The framework is general : no Markovian Assumption
- Can solve all the pre-existing models without ad-hoc methods
- How ?
  - Use a Dynamic Programming Approach (DPP)
- Why is it different from the literature :
  - Agent need to stochastic control problem for an arbitrary compensation scheme (possibly non-Markovian)
  - Principal need to optimize the contract for all possible (non-linear) reaction of the Agent.
    - Tools : calculus of variation, stochastic Pontryagin max. principle (Cvitanić and Zhang)
    - Ad-hoc (case-by-case basis) methods (cf. Holmström and Milgrom, Cvitanić, Possamaï, Touzi)

Introduction

- Dynamic programming ... seems simple no ?
  - Inspiration from Sannikov (2008)
- Restrict the family of admissible contracts to a collection that *can be solved using Dynamic Programming*
  - For this family, use standard verification methods
- However, this approach does not suffer from lack of generality
  - Under mild technical conditions, can express the Principal's optimum over this restricted collection *as equal* to the supremum *over all feasible contracts*.
    - Technical difficulties when Agent controls the diffusion terms
    - Can represen the Agent's value process as the solution of a BSDE
    - Even more : a 2BSDE, actually, as developed in Soner, Touzi, Zhang 2012.

Model and formalism – introduction

- The agent ('he') controls the evolution of a $d-$dimensional diffusion process X, with its effort $\nu = (\alpha, \beta)$
  - Through its drift $\lambda(\alpha)$
  - ... and the volatility $(\sigma(\beta))$!
- The principal ('she') does not observe the effort ν, but only the process X over time.
- She pays a compensation ξ (a contract) contingent on$X$ at terminal date T
- The agent chooses its effort maximizing its final utility $U_A(\xi)$, subject to some cost c_t and discounting k_t.
- The principal chooses the contract maximizing its utility $U_P(\ell(X) - \xi)$.

Formalism – control models

- The agent controls the SDE of the state variable (the *output* process)

$$X_t = X_0 + \int_0^t \sigma_s(X_\cdot, \beta_t)[\lambda_s(X_\cdot, \alpha_s)ds + dW_s]$$

- The couple $\mathbb{M} = (\mathbb{P}, \nu)$ is a *control model* if X^M is a *weak solution* of the controlled state equation.
  - *'Recall'* : A weak solution of a 'path-dependent' SDE is a tuple $(\Omega, \mathcal{F}, \mathbb{P}, W, X)$ such that $(\Omega, \mathcal{F}, \mathbb{P})$ is a proba space, (W, X) two stochastic processes, W a $(\mathcal{F}^W, \mathbb{P})$-Brownian motion and the equation holds.
- We assume the set of control models is $\mathcal{M} \ni \mathbb{M}$ non-empty.

Formalism – Agent's problem

- A r.v. ξ is called a contract if it contingent on$X$ at terminal date T, (i.e. ξ is $\mathcal{F}_T$-measurable) and with some L^p-moments.
- Let c be cost function, assumed to have some measurability and L^p regularity for all effort $\mathbb{M} \in \mathcal{M}$
- Let $\mathcal{K}_t = \exp(-\int_0^t k_s(\nu_s)ds)$ be a discount factor, with k_t bounded and optional.
- The Agent will aim at maximizing an objective function :

$$J^A(\mathbb{M}, \xi) := \mathbb{E}^{\mathbb{P}}\left[\mathcal{K}_T \xi - \int_0^T \mathcal{K}_t c_t(\nu_t)dt\right]$$

- The optimal effort will be to choose the best control model $(\mathbb{P}^\star, \nu^\star) \in \mathcal{M}^\star(\xi)$ for a given contract :

$$V^A(\xi) := \sup_{\mathbb{M} \in \mathcal{M}} J^A(\mathbb{M}, \xi)$$

Formalism – Agent's problem – Remarks

- In the previous slide, the agent was risk-neutral. However, one can replace ξ by a utility function U^A :

$$J^A(\mathbb{M},\xi) := \mathbb{E}^{\mathbb{P}}\left[\mathcal{K}_T U_A(\xi) - \int_0^T \mathcal{K}_t c_t(\nu_t)dt\right]$$

- The utility is separable btw the compensation ξ and the cost c_t.
- One could also consider the objective as :

$$J^A(\mathbb{M},\xi) := \mathbb{E}^{\mathbb{P}}\left[\exp\Big(- \mathrm{sgn}(U_A)\int_0^T \mathcal{K}_t c_t(\nu_t)\Big)\mathcal{K}_T U_A(\xi)\right]$$

  - In the following, to adapt for such a extension, one will need to replace ξ in the principal problem by $(U_A)^{-1}(\xi)$
  - Alternatively, one can think about ξ as compensation in 'utility'.
  - Recall that $V^A(\xi) := \sup_{\mathbb{M}\in\mathcal{M}} J^A(\mathbb{M},\xi)$ is the 'value function'.

Formalism – Principal's problem

- The principal will choose a contract which is *admissible* i.e. $\xi \in \Xi$

$$\Xi := \{\xi \in \mathcal{C}_0, \mathcal{M}^\star(\xi) \neq \emptyset, \text{and} V^A(\xi) \geq R\}$$

  where R is the reservation utility of the agent.

- Let $\ell(X)$ be liquidation value, and $\mathcal{K}_t^P = \exp(-\int_0^t k_s^P(\nu_s)ds)$ be a discount factor, with k_t bounded and optional.

$$J^P(\xi) = \sup_{(\mathbb{P}^\star,\nu^\star)\in\mathcal{M}^\star} \mathbb{E}^{\mathbb{P}}\left[\mathcal{K}_t^P U(\ell - \xi)\right]$$

- The value function defines :

$$V^P := \sup_{\xi\in\Xi} J^P(\xi)$$

Formalism – Comments

1. The problems are non-standard : $\xi|\mathcal{F}_t$ can be Non-Markovian and thus the Dynamic Programming Principle (DPP) would not be valid for both the agent and the principal.
   - The main goal of this article is to reduce these problems to those that can be solved using DPP.

Formalism – Comments

1. The problems are non-standard : $\xi|\mathcal{F}_t$ can be Non-Markovian and thus the Dynamic Programming Principle (DPP) would not be valid for both the agent and the principal.
    - The main goal of this article is to reduce these problems to those that can be solved using DPP.

2. The weak-formulation of the SDE is standard in continuous-time Principal Agent models : the agent's efforts ν affect the output thought the distribution $\mathbb{P}$. Moreover, Principal's contract will only be $\sigma(X_t)-$adapted and so will be her information.
    - This difference highlight the difference in information between the Principal and the Agent.

A restricted class of contract

- The idea being to solve the problem with dynamic programming (DPP), we now focus on a solution methods 'as if' it was possible to use DPP.
- The main theorem of the paper shows that the optimal contracts in this class indeed reaches the same value as the restricted
- In the following, I describe the family of restricted contracts :
  - *'Recall'* : The 'standard' approach from stochastic control [the verification method] consists in solving a HJB [Hamilton-Jacobi-Bellman] equation, finding the optimal feedback control and verifying that the underlying stochastic process solves the SDE.
  - The heuristic derivation of the HJB is detailed here.

Restricted class of contract – The HJB equation

- The Hamiltonian of the problem considered above is the following :

$$H_t(x, y, z, \gamma) = \sup_{u \in A \times B} h_t(x, y, z, \gamma, u)$$

$$h_t(x, y, z, \gamma, u) = -c_t(x, u) - k_t(x, u)\, y + \sigma_t(x, b)\lambda_t(x, a) \cdot z + {}^1\!/_2\, \mathrm{Tr}(\sigma_t \sigma_t^T \gamma)$$

  Suppose :
  - If the coeff λ, σ, c, k are not path dependent, i.e. depend on x only through the current value x_t
  - The contract ξ depends on x only through the final value x_T
- *then*, by verification theorem, the Agent's value function is $V^A(\xi) = v(0, X_0)$ where $v(t, x)$ is the unique viscosity solution of the HJB :

$$-\partial_t v(t, x) - H_t(x, v, Dv, D^2 v) = 0, \qquad v(T, x) = g(x), \quad \forall (t, x) \in [0, T) \times \mathbb{R}^d$$

Restricted class of contract – The HJB equation

- In the Markovian setting described before, assuming v solution of the HJB is $\mathcal{C}^{1,2}$ we can introduce the $V_t(\xi) = v(t, x_t)$
- Therefore, by definition of the value function we have $v(T, x_T) = g(x_T) = \xi(x_T)$
- The optimal compensation ξ being simply the value function v, we can obtain the following representation, by the Itô's formula :

$$g(X_T) = v(0, X_0) + \int_0^T z_t \cdot dX_t + \int_0^T \tfrac{1}{2}\mathrm{Tr}(\gamma_t \, d\langle X\rangle_t) - H_t(V_t, z_t, \gamma_t)dt$$

with $V_t = v(t, x_t),\ z_t = Dv(t, x_t),\ \gamma_t = D^2 v(t, x_t)$

- This formulation for optimal contract is inspired from Sannikov.
- The main idea will thus be to express V_t in term of ξ, i.e. a BSDE formulation !

Restricted class of contract – Definition

The collection $\mathcal{V}$ of predictable process (Z, Γ) is *defined* such that :

- The process $Y^{Z,\Gamma}$ and Z have some L^p regularity/integrability :

$$Y^{Z,\Gamma} := Y_0 + \int_0^t Z_s \cdot dX_s + \int_0^t \tfrac{1}{2}\mathrm{Tr}(\Gamma_s \, d\langle X\rangle_s) - H_s(V_s, Z_s, \Gamma_s) ds$$

  - This process will be central, as representation of Agent's value fct for the Principal.

- There exists a (weak-)solution $(\mathbb{P}^{Z,\Gamma}, \nu^{Z,\Gamma}) \in \mathcal{M}$ maximizing the hamiltonian :

$$H_t(X_t, Y_t, Z_t, \Gamma_t) = h_t(X_t, Y_t, Z_t, \Gamma_t, \nu_t^{Z,\Gamma}) \qquad \mathbb{P}^{Z,\Gamma} - a.e$$

  - It is, in a way, the idea of finding an optimal feedback control in the verification approach (given v, i.e. Y here).

Restricted class of contract – A verification argument

Prop. 3.3 is an important result, used in the proof of the main theorems :
For $Y_0 \in \mathbb{R}$ and $(Z, \Gamma) \in \mathcal{V}$ we have :

- $Y_T^{Z,\Gamma} \in \mathcal{C}_0$
  - The terminal value Y will be a suitable contract
- $Y_0 = V^A(Y_T^{Z,\Gamma})$ and any couple $(\mathbb{P}^{Z,\Gamma}, \nu^{Z,\Gamma})$ will be an optimal response to such contract, i.e. $(\mathbb{P}^{Z,\Gamma}, \nu^{Z,\Gamma}) \in \mathcal{M}^\star(Y_T^{Z,\Gamma})$
  - For such type of contracts, agent's value coincide with $Y_t^{Z,\Gamma}$.
- $(\mathbb{P}^\star, \nu^\star) \in \mathcal{M}^\star(Y_T^{Z,\Gamma})$ *if and only if*

$$H_t(X_t, Y_t, Z_t, \Gamma_t) = h_t(X_t, Y_t, Z_t, \Gamma_t, \nu_t^\star) \quad \mathbb{P}^\star - a.e$$

  - Optimal actions $\nu^\star$ coincide/ are identified with hamiltonian maximizers (on the support of $\mathbb{P}^\star$).

Restricted class of contract – A verification argument

Prop. 3.3, Ideas of the proof : For $Y_0 \in \mathbb{R}$ and $(Z, \Gamma) \in \mathcal{V}$ we have :

- $Y_0 = V^A(Y_T^{Z,\Gamma})$ and any couple $(\mathbb{P}^{Z,\Gamma}, \nu^{Z,\Gamma})$ will be an optimal response to such contract, i.e. $(\mathbb{P}^{Z,\Gamma}, \nu^{Z,\Gamma}) \in \mathcal{M}^\star(Y_T^{Z,\Gamma})$
- $(\mathbb{P}^\star, \nu^\star) \in \mathcal{M}^\star(Y_T^{Z,\Gamma})$ *if and only if*

$$H_t(X_t, Y_t, Z_t, \Gamma_t) = h_t(X_t, Y_t, Z_t, \Gamma_t, \nu_t^\star) \quad \mathbb{P}^\star - a.e$$

Restricted class of contract – Notations

- Since we have identified the optimal effort in such setting, we denote them $u^\star = (\alpha^\star, \beta^\star)$:

$$H_t(x, y, z, \gamma_t) = h_t(y, z, \gamma_t, \nu_t^\star)$$

- The optimal feedback control induces drift and variance :

$$\lambda_t^\star(x, y, z, \gamma) = \lambda_t(x, \alpha_t^\star(x, y, z, \gamma)) \quad \text{and} \quad \sigma_t^\star(x, y, z, \gamma) = \sigma_t(x, \beta_t^\star(x, y, z, \gamma))$$

- The output process rewrites :

$$X_t = X_0 + \int_0^t \sigma_t^\star(X, Y_s, Z_s, \Gamma_s)\big[\,\lambda^\star(X, Y_s, Z_s, \Gamma_s)ds + dW_s\big], \quad \forall t \in [0, T]$$

  - Note that for $\lambda^\star, \sigma^\star$ given, the SDE is *controlled* by (z, γ)

Restricted class of contract – Principal's point of view

- The previous verification argument allows to determine the 'agent-optimal' contract as the value function of the Agent.
- The authors show – and that the *main result* of the article – that it correspond to the optimum for the Principal problem
- Informally, it will means to prove that

$$V^P := \sup_{\xi \in \Xi} J^P(\xi) = \sup_{\substack{\xi^\star \equiv Y_T^{Z,\Gamma}, \\ Y_0 \geq R,\, (Z,\Gamma) \in \mathcal{V}}} \underline{V}(Y_0)$$

'heuristically', and where $\underline{V}(Y_0)$ remains to define.

Restricted class of contract – Principal's point of view

Prop. 3.4, a direct consequence of prop 3.3.

- The principal's value function is minored by the maximum over restricted contract :
- Defining

$$\underline{V}(Y_0) := \sup_{(Z,\Gamma)\in\mathcal{V}} \sup_{(\mathbb{P},\nu)\in\mathcal{M}^\star} \mathbb{E}^{\mathbb{P}}\left[\mathcal{K}_t^P U(\ell - Y_T^{Z,\Gamma})\right]$$

- We have (Prop 3.4) :

$$V^P := \sup_{\xi\in\Xi} J^P(\xi) \geq \sup_{Y_0\geq R} \underline{V}(Y_0)$$

- Intuitively, the RHS implies to choose an optimal contract s.t. :
  - (i) initial value Y_0 is above reservation utility
  - (ii) agent's value fct will coincide with $(Y_t^{Z,\Gamma})_t$ (resp. cond. of $\mathcal{V}$)
  - (iii) the agent will behave optimally to the contract given by $Y_T^{Z,\Gamma}$

Restricted class of contract – Main reduction result

Theorem 3.6

- Assume that $\mathcal{V} \neq \emptyset$
- *then* we have :

$$V^P = \sup_{Y_0 \geq R} \underline{V}(Y_0)$$

- Moreover, the maximizer of LHS optim $(Y_0^\star, Z^\star, \Gamma^\star)$ induces an optimal contract $\xi^\star := Y_T^{Z^\star, \Gamma^\star}$.
  - Since the LHS happens to be the value function of a standard (DPP-style) stochastic control problem,
  - The assumption $\mathcal{V} \neq \emptyset$ is mild (for $\underline{V} \neq -\infty$).
- Before presenting the sketch of the proof in a specific case, I derive the solution of Principal's control pblm

Restricted class of contract – Solving Principal's HJB

- Assuming $\mathcal{M}^\star \neq \emptyset$

$$\underline{V}(Y_0) := \sup_{(Z,\Gamma)\in\mathcal{V}} \sup_{(\mathbb{P},\nu)\in\mathcal{M}^\star} \mathbb{E}^{\mathbb{P}}\left[\mathcal{K}_t^P U(\ell - Y_T^{Z,\Gamma})\right]$$

- It is a "standard" problem to solve
  - It correspond to the controlled SDE :

$$dY_t^{Z,\Gamma} = \left(Z_t \cdot \sigma_t^\star \lambda_t^\star + \frac{1}{2}\mathrm{Tr}(\sigma_t^\star \sigma_t^{\star\,T}\Gamma_t) - H\right)(Y_t^{Z,\Gamma}, Z_t, \Gamma_t)dt + Z_t \cdot \sigma_t^\star (Y_t^{Z,\Gamma}, Z_t, \Gamma_t) dW_t^{\mathbb{M}^\star}$$

  - The (long) Hamiltonian :

$$G(t,x,y,p,M) := \sup_{(z,\gamma)} \sup_{u^\star} \Big\{ (\sigma_t^\star \lambda_t^\star)\cdot p_x + \left(z \cdot \sigma_t^\star \lambda_t^\star + \tfrac{1}{2}\mathrm{Tr}(\sigma_t^\star \sigma_t^{\star\,T}\gamma_t) - H_t\right)(x,y,z,\gamma)\, p_y$$
$$+ \frac{1}{2}\mathrm{Tr}\left(\sigma_t^\star \sigma_t^{\star\,T}(M_{xx} + zz^T M_{yy})\right) + \sigma_t^\star \sigma_t^{\star\,T}(x,y,z,\gamma)\, z \cdot M_{xy} \Big\}$$

Stochastic control – Solving Principal's HJB

- The (long) Hamiltonian of Principal's problem :

$$G(t,x,y,p,M) := \sup_{(z,\gamma)} \sup_{u^\star} \Big\{ (\sigma_t^\star \lambda_t^\star) \cdot p_x + \big(z \cdot \sigma_t^\star \lambda_t^\star + \tfrac{1}{2} \mathrm{Tr}(\sigma_t^\star {\sigma_t^\star}^T \gamma_t) - H_t \big)(x,y,z,\gamma)\, p_y$$
$$+ \frac{1}{2} \mathrm{Tr}\big(\sigma_t^\star {\sigma_t^\star}^T (M_{xx} + zz^T M_{yy}) \big) + \sigma_t^\star {\sigma_t^\star}^T (x,y,z,\gamma)\, z \cdot M_{xy} \Big\}$$

with $M =: \begin{pmatrix} M_{xx} & M_{xy} \\ M_{yx} & M_{yy} \end{pmatrix} \in \mathcal{S}_{d+1}$ and $p =: \begin{pmatrix} p_x \\ p_y \end{pmatrix} \in \mathbb{R}^{d+1}$ Comments :

- The maximizat° of the Hamiltonian is made over $(z,\gamma) \in \mathbb{R} \times \mathcal{S}_d(\mathbb{R})$ and $u^\star = (\alpha^\star, \beta^\star)$ implies the drift/diffusion terms $\lambda^\star$ and $\sigma^\star$.
- Assume the existence of $(\hat{z},\hat{\gamma})(t,x,y,p,M)$ maximizer of the Hamiltonian
- The value function also depends on y which is the value fct of the agent.

Stochastic control – Solving Principal's HJB

- Let $v \in \mathcal{C}^{1,2}([0,T), \mathbb{R}^{n+1}) \cap \mathcal{C}^0([0,T] \times \mathbb{R}^{d+1})$ a classical solution of the HJB :

$$\begin{cases} (\partial_t v - k^P)(t,x,y) \ + \ G(t,x,y,Dv,D^2v) = 0 \quad \forall (t,x,y) \in [0,T) \times \mathbb{R}^d \times \mathbb{R} \\ v(T,x,y) = U(\ell(x) - y) \end{cases}$$

- Assuming that :
  - $v(t, X_t, Y_t)_t$ is U.I (uniform integrable) $\forall (\mathbb{P}, \nu) \in \mathcal{M}^\star, \ \forall (Z, \Gamma) \in \mathcal{V}$
  - The Hamiltonian has maximizers $(\hat{z}, \hat{\gamma})$ s.t.
    - The controlled SDE governing X_t and$Y_t^{Z,\Gamma}$ with controls $(Z^\star, \Gamma^\star) = (\hat{z}, \hat{\gamma})(\cdot, Dv, D^2v)(t, X_t, Y_t)$ has a weak solution $(\mathbb{P}^\star, \nu^\star)$
    - $(Z^\star, \Gamma^\star) \in \mathcal{V}$.
- *Then* $\underline{V}(Y_0) = v(0, X_0, Y_0)$ and $(Z^\star, \Gamma^\star)$ is an optimal control for Principal's problem.

Forking

- A concrete example, from Cvitanić, Wan and Zhang (2009)
- The proof in the special case where the Agent does not control the volatility of output.
  - Remember, the idea is to proove that $V^P := \sup_{\xi \in \Xi} J^P(\xi) = \sup_{\substack{\xi^\star \equiv Y_T^{Z,\Gamma}, \\ Y_0 \geq R,\ (Z,\Gamma) \in \mathcal{V}}} \underline{V}(Y_0)$
  - But, the class Ξ is only $\mathcal{F}_T$-mesurable : not possible to use controlled SDE (and DPP-style stuffs)
  - Instead, characterize the process $Y_t^{Z,\Gamma}$ as a (controlled) BSDE
  - Prove the existence/uniqueness result from the famous results of *Pardoux and Peng 90*
- The proof in the general case is similar :
  - In this case, trouble comes from the 2nd order, diffusion term
  - Instead, characterize the process $Y_t^{Z,\Gamma}$ as a (controlled) *Second-Order* BSDE
  - Use results from Soner, Touzi and Zhang (2012)

Fixed volatility of output

- Suppose the agent has no action on volatility :
  - The agent's hamiltonian reduces to

$$H(x,y,z,\gamma) = \frac{1}{2}\mathrm{Tr}(\sigma_t\sigma_t^T\gamma) + F_t(x,y,z)$$

where $$F_t(x,y,z,a) = \sup_{a\in A}\left\{-c_t(x,a)-k_t(x,a)y+\sigma_t(x)\lambda_t(x,a)\cdot z\right\}$$

- The dynamics of the reduced contract becomes :

$$Y_t^Z := Y_0 + \int_0^t Z_s \cdot dX_s - \int_0^t F_t(X, Y_s^Z, Z_s)ds$$

  - To be able to use the *Thm 3.6*, we need to represent any contract $\xi \in \Xi$ as a compensation of the form $\xi = Y_T^Z$
  - It reduces the problem to solving a BSDE :

$$Y_0 = \xi + \int_0^T F_t(X, Y_s^Z, Z_s)ds - \int_0^T Z_s \cdot dX_s$$

Fixed volatility of output

- The process $Y_t^{Z,\Gamma}$, because it depends on the contract $\xi = Y_T^{Z,\Gamma}$ is a typical example of Backward Stochastic differential equation.
- Starting from :

$$\begin{cases} Y_t^Z = Y_0 - \int_0^t F(X, Y_s^Z, Z_s) ds + \int_0^t Z_s \cdot dX_s \\ Y_T^Z = \xi \end{cases}$$

In the following : BSDE - definitions and main results

Fixed volatility of output

- *'Recall'* that the predictable representation property of a semi martingale X w.r.t./under $(\mathbb{F}, \mathbb{Q})$ if any $(\mathbb{F}, \mathbb{Q})-$local-martingale Y can be written in the form $Y_t = m + \int_0^t Z_s dX_s$ where Z_t is a predictable process and m a $\mathcal{F}_0$-measurable r.v.
- *'Recall'* the Blumenthal zero-one law : If

$$\mathcal{F}_{0+} = \bigcap_{u>0} \mathcal{F}_u$$

  then $\mathcal{F}_{0+}$ is trivial in the sense than $\forall A \in \mathcal{F}_{0+}, \mathbb{P}(A) = 0$ or 1
- According to the authors, the standard theory of BSDE directly implies that these two conditions, added to the standard regularity/integrability assumption + generator of the BSDE being uniform Lipschitz directly implies existence and uniqueness.

Some results on BSDE

- *'Recall'* : A fundamental result from Pardoux and Peng on BSDE :
  - A solution of the BSDE ...

$$\begin{cases} dY_t = -f(Y_t, Z_t)dt + Z_t dW_t \\ Y_T = \xi \qquad\qquad \in [0, T] \times \mathbb{R}^d \end{cases}$$

  - ... is a *couple* (Y_t, Z_t) satisfying some measurability/integrability conditions such that

$$Y_t = \xi + \int_t^T f_s(Y_s, Z_s)ds - \int_t^T Z_s dW_s$$

- Pardoux and Peng 90 : Existence and Unicity of solution of BSDE :
  - Assuming that f is uniformly Lipschitz in (y, z) and $\xi, f_t(0, 0)$ are L^2
  - *then* there exists a unique solution (Y, Z) to the BSDE.

The stochastic control problem – the HJB equation

- The aim of the agent is to maximize its objective function :

$$v(t_0, X_{t_0}) = \sup_{\{\alpha_t\}_{t_0}^T} \mathbb{E}_{t_0}\Big(\int_{t_0}^T L(t, X_t, \alpha_t)dt + g(X_T)\Big)$$

  where v is the value function of the agent (at time t_0), L and G resp. the running gain and terminal gain.

The stochastic control problem – the HJB equation

- The aim of the agent is to maximize its objective function :

$$v(t_0, X_{t_0}) = \sup_{\{\alpha_t\}_{t_0}^T} \mathbb{E}_{t_0}\Big(\int_{t_0}^T L(t, X_t, \alpha_t)dt + g(X_T) \Big)$$

  where v is the value function of the agent (at time t_0), L and G resp. the running gain and terminal gain.

- α_t the (adapted) control variable and X_t is the state variable, (unique) solution of SDE :

$$\begin{cases} dX_t = b(t, X_t, \alpha_t)dt + \sigma(t, X_t, \alpha_t)dB_t \\ X_{t_0} = x_0 \qquad\qquad (t_0, x_0) \in [0, T] \times \mathbb{R}^d \end{cases}$$

  where b is the drift, σ the variance and B_t a Brownian motion More on this.

The stochastic control problem – the HJB equation

- Here, Bellman dynamic programming principle holds :

$$v(t_0, X_{t_0}) = \sup_{\{\alpha_t\}_{t_0}^{T}} \mathbb{E}_{t_0}\Big(\int_{t_0}^{t_1} L(t, X_t, \alpha_t)dt + v(t_1, X_{t_1}) \Big)$$

The stochastic control problem – the HJB equation

- Here, Bellman dynamic programming principle holds :

$$v(t_0, X_{t_0}) = \sup_{\{\alpha_t\}_{t_0}^T} \mathbb{E}_{t_0}\Big(\int_{t_0}^{t_1} L(t, X_t, \alpha_t)dt + v(t_1, X_{t_1})\Big)$$

- The idea is to study ”infinitesimal” variation in the value function

The stochastic control problem – the HJB equation

- Here, Bellman dynamic programming principle holds :

$$v(t_0, X_{t_0}) = \sup_{\{\alpha_t\}_{t_0}^T} \mathbb{E}_{t_0}\Big(\int_{t_0}^{t_1} L(t, X_t, \alpha_t)dt + v(t_1, X_{t_1}) \Big)$$

- The idea is to study "infinitesimal" variation in the value function
- Use the Itô formula here to compute the value fct at time $t + h$:

$$\sup_{\{\alpha_t\}} \mathbb{E}_{t_0}\Big(\int_{t_0}^{t_0+h} L(t, x, \alpha_t)dt + \int_{t_0}^{t_0+h} \Big\{ \partial_t v + \nabla_x v \cdot \boldsymbol{b}_t + \frac{1}{2} Tr\big(\sigma_t \sigma_t^T D_{xx}^2 v\big) \Big\} dt + \int_{t_0}^{t_0+h} \nabla_x v \cdot \sigma_t \, d\mathbf{B}_t \Big) = 0$$

The stochastic control problem – the HJB equation

- Here, Bellman dynamic programming principle holds :

$$v(t_0, X_{t_0}) = \sup_{\{\alpha_t\}_{t_0}^{T}} \mathbb{E}_{t_0}\Big(\int_{t_0}^{t_1} L(t, X_t, \alpha_t)dt + v(t_1, X_{t_1})\Big)$$

- The idea is to study ”infinitesimal” variation in the value function
- Use the Itô formula here to compute the value fct at time $t+h$:

$$\sup_{\{\alpha_t\}} \mathbb{E}_{t_0}\Big(\int_{t_0}^{t_0+h} L(t,x,\alpha_t)dt + \int_{t_0}^{t_0+h} \Big\{ \partial_t v + \nabla_x v \cdot \boldsymbol{b}_t + \frac{1}{2} Tr\big(\sigma_t \sigma_t^T D_{xx}^2 v\big) \Big\} dt + \int_{t_0}^{t_0+h} \nabla_x v \cdot \sigma_t \, d\mathbf{B}_t \Big) = 0$$

- The expectation of stochastic integral is zero $\mathbb{E}(\int_0^t \cdots dB_s) = B_0 = 0$ by martingale property.

The stochastic control problem – the HJB equation

- Here, Bellman dynamic programming principle holds :

$$v(t_0, X_{t_0}) = \sup_{\{\alpha_t\}_{t_0}^T} \mathbb{E}_{t_0}\Big(\int_{t_0}^{t_1} L(t, X_t, \alpha_t)dt + v(t_1, X_{t_1})\Big)$$

- The idea is to study "infinitesimal" variation in the value function
- Use the Itô formula here to compute the value fct at time $t+h$:

$$\sup_{\{\alpha_t\}} \mathbb{E}_{t_0}\Big(\int_{t_0}^{t_0+h} L(t,x,\alpha_t)dt + \int_{t_0}^{t_0+h} \Big\{\partial_t v + \nabla_x v \cdot \boldsymbol{b}_t + \frac{1}{2} Tr\big(\sigma_t \sigma_t^T D_{xx}^2 v\big)\Big\} dt + \int_{t_0}^{t_0+h} \nabla_x v \cdot \sigma_t \, d\mathbf{B}_t \Big) = 0$$

- The expectation of stochastic integral is zero $\mathbb{E}(\int_0^t \cdots dB_s) = B_0 = 0$ by martingale property.
- Take $h \to 0$, the integrand need to be zero for every t :

$$\partial_t v(t,x) + \sup_a \Big\{ L(t,x,a) + \nabla_x v(t,x) \cdot b(t,x,a) + \frac{1}{2} Tr\big(\sigma(t,x,a)\sigma(t,x,a)^T D_{xx}^2 v(t,x)\big)\Big\} = 0$$

The stochastic control problem – the HJB equation

- Here, Bellman dynamic programming principle holds :

$$v(t_0, X_{t_0}) = \sup_{\{\alpha_t\}_{t_0}^{T}} \mathbb{E}_{t_0}\Big(\int_{t_0}^{t_1} L(t, X_t, \alpha_t)dt + v(t_1, X_{t_1})\Big)$$

- The idea is to study ”infinitesimal” variation in the value function
- Use the Itô formula here to compute the value fct at time $t+h$:

$$\sup_{\{\alpha_t\}} \mathbb{E}_{t_0}\Big(\int_{t_0}^{t_0+h} L(t,x,\alpha_t)dt + \int_{t_0}^{t_0+h}\Big\{\partial_t v + \nabla_x v \cdot \boldsymbol{b}_t + \frac{1}{2}Tr\big(\sigma_t\sigma_t^T D_{xx}^2 v\big)\Big\}dt + \int_{t_0}^{t_0+h} \nabla_x v \cdot \sigma_t \, d\mathbf{B}_t\Big) = 0$$

- The expectation of stochastic integral is zero $\mathbb{E}(\int_0^t \cdots dB_s) = B_0 = 0$ by martingale property.
- Take $h \to 0$, the integrand need to be zero for every t :

$$\partial_t v(t,x) + \sup_a \Big\{ L(t,x,a) + \nabla_x v(t,x) \cdot b(t,x,a) + \frac{1}{2} Tr\Big(\sigma(t,x,a)\sigma(t,x,a)^T D_{xx}^2 v(t,x)\Big)\Big\} = 0$$

- This is the Hamilton Jacobi Bellman (HJB) PDE !

The stochastic control problem – the HJB equation

- The Hamilton-Jacobi-Bellman :

$$\partial_t v(t,x) + \sup_a \left\{ L(t,x,a) + \nabla_x v(t,x) \cdot b + \frac{1}{2} Tr\left(\sigma\sigma^T D^2_{xx} v(t,x)\right) \right\} = 0$$

- Or writing it with "Hamiltonians"

$$H(t,x,p,M) = \sup_a \left\{ L(t,x,a) + p \cdot b + \frac{1}{2} Tr\left(\sigma\sigma^T M\right) \right\} = 0$$

- the HJB rewrites :

$$\partial_t v(t,x) + H(t,x,\nabla_x v, D^2_{xx} v) = 0$$

- The optimal control can be given in feedback form by the First-Order Conditions (FOC).

The stochastic control problem – Solutions

- Verification approach (the 'standard' approach of stochastic control) :
  - Find $w(t,x)$ a solution of the HJB equation.
  - Find a mesurable fct $a(t,x)$ maximizing the hamiltonian (for this w).
  - Plug the $a(t,X_t)$ is the dynamics $dX_t = b(\cdot)dt + \sigma(\cdot)dW_t$.
  - If this SDE has a solution $\hat{X}_t^a$ given initial condition (t,x),
- *then* : the function w is the value function of the stochastic control problem.
  - What if the fct v is not smooth ? (not $\mathcal{C}^{1,2}$)
    $\rightarrow$ Viscosity solutions : Crandall and Lions (1989)

Rappels : Itô's formula

- For any X_t Itô process :

$$dX_t = b_t\,dt + \sigma_t\,dB_t$$

and any $\mathcal{C}^{1,2}$ scalar function $f(t,x)$ of two real variables t and x, one has :

$$df(t,X_t) = \left(\frac{\partial f}{\partial t} + b_t\frac{\partial f}{\partial x} + \frac{\sigma_t^2}{2}\frac{\partial^2 f}{\partial x^2}\right)dt + \sigma_t\frac{\partial f}{\partial x}\,dB_t$$

- For vector-valued processes $\mathbf{X}_t = (X_t^1, X_t^2, \dots, X_t^n)$

$$d\mathbf{X}_t = \boldsymbol{b}_t\,dt + \sigma_t\,d\mathbf{B}_t$$

- The Itô formula rewrites :

$$\begin{aligned} df(t,\mathbf{X}_t) &= \frac{\partial f}{\partial t}(t,X_t)\,dt + \sum_{i=1}^{d}\frac{\partial f}{\partial x_i}(t,X_t)dX_t^i + \frac{1}{2}\sum_{i,j=1}^{d}\frac{\partial^2 f}{\partial x_i\partial x_j}(t,X_t)d\langle X^i,X^j\rangle_t \\ &= \partial_t f\,dt + \nabla_x f\cdot d\mathbf{X}_t + \frac{1}{2}Tr\left(\sigma_t\sigma_t^T D_{xx}^2 f\right)\,dt, \\ &= \left\{\partial_t f + \nabla_x f\cdot\boldsymbol{b}_t + \frac{1}{2}Tr\left(\sigma_t\sigma_t^T D_{xx}^2 f\right)\right\}dt + \nabla_x f\cdot\sigma_t\,d\mathbf{B}_t \end{aligned}$$

Go back