Notes on Stochastic Control

Haosheng Zhou

Dec, 2022

Contents

The following contents about HJE & HJBE refers to the Evans book.

Hamilton-Jacobi Equation (HJE)

The Hamilton-Jacobi equation is a non-linear first-order PDE with the form

$$\begin{cases} u_t + H(Du) = 0 & in\ \mathbb{R}^n \times (0,\infty) \\ u = g & on\ \mathbb{R}^n \times \{t=0\} \end{cases} \tag{1}$$

where $u = u(x,t) : \mathbb{R}^n \times [0,\infty) \to \mathbb{R}$ is the function to solve out and $Du = (u_{x_1}, ..., u_{x_n})$ is the gradient of u w.r.t. space variable x. Here the **Hamiltonian** $H : \mathbb{R}^n \to \mathbb{R}$ is given and the initial condition $g : \mathbb{R}^n \to \mathbb{R}$ is given.

Connection with Hamilton's Equations

Let's first apply the method of characteristics to get some intuition by noticing that this equation is a first-order equation. We know that the method of characteristics does not necessarily hold in general (since it requires the existence of C^2 solution), but this may tell us how to proceed. In this section, we assume that HJE looks like

$$\begin{cases} u_t + H(Du, x) = 0 & in\ \mathbb{R}^n \times (0,\infty) \\ u = g & on\ \mathbb{R}^n \times \{t=0\} \end{cases} \tag{2}$$

where the Hamiltonian also depends on x.

Notice that here we **merge the time variable t with the space variable x and denote it as $x \in \mathbb{R}^{n+1}$, where $x^1, ..., x^n$ are components of x and x^{n+1} denotes the time**. Define

$$z(s) = u(x(s)) \tag{3}$$

as the version of u along the characteristic curve and

$$p(s) = Du(x(s)) \in \mathbb{R}^{n+1} \tag{4}$$

as the version of Du along the characteristic curve, note that here $p^1, ..., p^n$ are partial derivatives w.r.t. x-components and p_{n+1} is the partial derivative w.r.t. t. One would always set the characteristic direction to be

$$x'(s) = D_p F \tag{5}$$

where the original PDE can be written as $F(Du, u, x) = 0$ and in this case

$$F(p, z, y) = p^{n+1} + H(p^1, ..., p^n, x^1, ..., x^n) \tag{6}$$

As a result, one get the ODE system from the method of characteristics

$$\begin{cases}[x^i(s)]' = H_{p_i}(p^1, ..., p^n, x^1, ..., x^n) \ (i = 1, 2, ..., n) \\ [x^{n+1}(s)]' = 1\end{cases} \tag{7}$$

so one can identify $x^{n+1}(s)$ as s, meaning that the parameter s is the same as the time variable $t = x^{n+1}$. The equation for $z(s)$ is $z'(s) = D_pF \cdot p(s)$, so

$$\begin{aligned}z'(s) &= \sum_{i=1}^{n} H_{p_i}(p^1, ..., p^n, x^1, ..., x^n) \cdot p^i(s) + p^{n+1}(s) \qquad (8)\\ &= \sum_{i=1}^{n} H_{p_i}(p^1, ..., p^n, x^1, ..., x^n) \cdot p^i(s) - H(p^1, ..., p^n, x^1, ..., x^n) \qquad (9)\end{aligned}$$

The equation for $p(s)$ is $p'(s) = -D_xF - D_zF \cdot p(s)$, so

$$\begin{cases}[p^i(s)]' = -H_{x_i}(p^1, ..., p^n, x^1, ..., x^n) \ (i = 1, 2, ..., n) \\ [p^{n+1}(s)]' = 0\end{cases} \tag{10}$$

with the last equation $[p^{n+1}(s)]' = 0$ as the redundant one since x^{n+1} has already been parameterized as s.

By cancelling all redundant equations and reorganizing the variables, we get the **characteristic ODE system** for HJE

$$\begin{cases}x'(s) = D_pH(p(s), x(s)) \\ z'(s) = D_pH(p(s), x(s)) \cdot p(s) - H(p(s), x(s)) \\ p'(s) = -D_xH(p(s), x(s))\end{cases} \tag{11}$$

where $p(s) = (p^1(s), ..., p^n(s))$ and $x(s) = (x^1(s), ..., x^n(s))$ (the last component in $x(s), p(s)$ is ignored). **The Hamilton's equation** is defined as the system consisting of the first and third equation, i.e.

$$\begin{cases}x'(s) = D_pH(p(s), x(s)) \\ p'(s) = -D_xH(p(s), x(s))\end{cases} \tag{12}$$

Remark. *The reason that we only take the equations w.r.t $x(s)$ and $p(s)$ in the Hamilton's equations is that those two equations have nothing to do with z, they already have $2n$ unknowns and $2n$ equations. In other words, the equation w.r.t. $z(s)$ does not provide any effective information for the derivation of $x(s), p(s)$, and after solving out $x(s), p(s)$, one can immediately know $z(s)$.*

A Problem in the Calculus of Variation

The connection between HJE and Hamilton's equations can also be shown in another perspective by considering a problem in the calculus of variation. The problem is formed as finding a best curve in an admissible class. The **admissible class** is defined as

$$\mathscr{A} = \left\{w \in C^2, w : [0,t] \to \mathbb{R}^n : w(0) = y, w(t) = x\right\} \tag{13}$$

so any admissible curve is a C^2 path in $\mathbb{R}^n$ such that it starts from point y and ends at point x with $x, y \in \mathbb{R}^n, t > 0$ given. Imagine $w(s) \in \mathscr{A}$ as the moving trajectory of a particle, then $w'(s)$ is actually the speed of the particle at each time. The **action functional** is then defined as

$$I[w] = \int_0^t L(w'(s), w(s))\, ds \tag{14}$$

where $L : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is a given smooth function called **Lagrangian** and we hope to find a curve $x(s) \in \mathscr{A}$ such that the action functional is minimized

$$I[x] = \inf_{w(s)\in\mathscr{A}} I[w] \tag{15}$$

Remark. *The Lagrangian has the meaning as the kinetic energy minus the potential energy in physics which has the meaning of "increments of distance". Here among all possible and smooth enough curves between two fixed points, we want to find $x(s)$ such that it minimizes the integral of the Lagrangian along the path, equivalent to saying that the optimal path is the one that* ***takes the "shortest" path****. If one still finds it hard to understand, think about how light travels, it always travels in the path such that the distance it goes through is the shortest, a natural minimization of a "trivial" action functional.*

Let's **assume** that the Lagrangian is given by $L = L(v, x)$ $(v, x \in \mathbb{R}^n)$ for the convenience of notations and that **the inf of** $I[w]$ **can be achieved by some** $x(s) \in \mathscr{A}$ **as the optimal path**. To build up a PDE for $x(s)$, choose smooth $y : [0, t] \to \mathbb{R}^n$ with $y(s) = (y^1(s), ..., y^n(s))$ such that $y(0) = y(t) = 0$ and consider perturbing the optimal path $x(s)$ by a small multiple of $y(s)$ to get

$$w(s) = x(s) + \tau y(s) \ (\tau \in \mathbb{R}) \tag{16}$$

since $w(s) \in \mathscr{A}$, one immediately sees that

$$I[w] \geq I[x] \tag{17}$$

Consider the action functional of the perturbed path

$$i : \mathbb{R} \to \mathbb{R}, i(\tau) = I[x + \tau y] \tag{18}$$

it's easy to see that it has minimum at $\tau = 0$ (assume it's differentiable) with

$$i'(\tau) = \frac{d}{d\tau}\int_0^t L(x'(s)+\tau y'(s), x(s)+\tau y(s))\,ds \tag{19}$$

$$= \int_0^t y'(s)\cdot L_v(x'(s)+\tau y'(s), x(s)+\tau y(s)) + y(s)\cdot L_x(x'(s)+\tau y'(s), x(s)+\tau y(s))\,ds \tag{20}$$

so

$$i'(0) = \int_0^t y'(s)\cdot L_v(x'(s), x(s)) + y(s)\cdot L_x(x'(s), x(s))\,ds \tag{21}$$

$$= \int_0^t \sum_{i=1}^n \left([y^i(s)]'\cdot L_{v_i}(x'(s), x(s)) + y^i(s)\cdot L_{x_i}(x'(s), x(s))\right)\,ds \tag{22}$$

$$= 0 \tag{23}$$

Do transformations to this integral to find

$$\sum_{i=1}^n \int_0^t \left([y^i(s)]'\cdot L_{v_i}(x'(s), x(s)) + y^i(s)\cdot L_{x_i}(x'(s), x(s))\right)\,ds \tag{24}$$

$$= \sum_{i=1}^n \int_0^t L_{v_i}(x'(s), x(s))\,dy^i(s) + \int_0^t L_{x_i}(x'(s), x(s))\cdot y^i(s)\,ds \tag{25}$$

$$= \sum_{i=1}^n -\int_0^t y^i(s)\,dL_{v_i}(x'(s), x(s)) + \int_0^t L_{x_i}(x'(s), x(s))\cdot y^i(s)\,ds \tag{26}$$

$$= \sum_{i=1}^n \int_0^t \left[-\frac{d}{ds}L_{v_i}(x'(s), x(s)) + L_{x_i}(x'(s), x(s))\right] y^i(s)\,ds \tag{27}$$

$$= 0 \tag{28}$$

which is valid for any smooth y such that $y(0) = y(t) = 0$. By density argument,

$$\forall i = 1, 2, ..., n, \forall s\in[0,t], -\frac{d}{ds}L_{v_i}(x'(s), x(s)) + L_{x_i}(x'(s), x(s)) = 0 \tag{29}$$

Theorem 1. ***(Euler-Lagrange Equation)*** *If path $x(s)$ is the optimal path and solves the variational problem mentioned above, then it must satisfy Euler-Lagrange equation that*

$$\forall s\in[0,t], -\frac{d}{ds}D_vL(x'(s), x(s)) + D_xL(x'(s), x(s)) = 0 \tag{30}$$

Remark. *The Euler-Lagrange equation consists of n* ***second-order ODEs****. Note that when $x(s)$ is the solution to the Euler-Lagrange equation, it does not necessarily achieve the inf of the action functional in the variational problem so* ***the converse of this theorem is not true****.*

In order to link Euler-Lagrange equation back to Hamilton's equations, let's first define

$$p(s) = D_v L(x'(s), x(s)) \tag{31}$$

as the **generalized momentum for position $x(s)$ and velocity $x'(s)$** (we will see why this has something to do with momentum later). We have to **assume that given $x, p \in \mathbb{R}^n$ the equation $p = D_v L(v, x)$ can be uniquely solved for v as a smooth function of p and x as $v(p, x)$**. The **Hamiltonian H associated with Lagrangian L** is defined as

$$H(p, x) = p \cdot v(p, x) - L(v(p, x), x) \ (p, x \in \mathbb{R}^n) \tag{32}$$

for $v(p, x)$ satisfying $p = D_v L(v, x)$ for given x, p defined implicitly.

Remark. *To understand the motivation of those definitions, let's consider the classical setting in physics that*

$$L(v, x) = \frac{1}{2} m||v||_2^2 - \phi(x) \tag{33}$$

where $\frac{1}{2}m||v||_2^2$ is the kinetic energy and ϕ is the potential energy with the mass $m > 0$. The Lagrangian immediately tells us that the Euler-Lagrange equation is

$$-\frac{d}{ds} m x'(s) - D\phi(x(s)) = 0 \tag{34}$$

$$m \cdot x''(s) = -D\phi(x(s)) \tag{35}$$

*where $D\phi$ is the force field generated by ϕ and this is **Newton's second law** for the acceleration of a particle with mass m in such force field.*

Let's then try to calculate the generated momentum

$$p(s) = m \cdot x'(s) \tag{36}$$

which is consistent with the true momentum in this case. The implicit definition of v is then

$$p(s) = D_v L(v(s), x(s)) \tag{37}$$

$$m \cdot x'(s) = m \cdot v(s) \tag{38}$$

since it can be uniquely solved for v as a smooth function, it must be true that $v(p, x) = x'(s)$, just the velocity of the particle. As a result, the Hamiltonian for such Lagrangian is

$$H(p, x) = m \cdot v \cdot v - L(v, x) \tag{39}$$

$$= \frac{1}{2} m||v||_2^2 + \phi(x) \tag{40}$$

so the Hamiltonian is the ***total energy*** *as the sum of kinetic and potential energy.*

Remark. *Another understanding of the definition of Hamiltonian is to specify p as the dual variable and L as the running cost in the minimization problem. Although it's currently not obvious where those names come from, this understanding is the closest to the stochastic control problem and the Pontryagin maximum principle we will discuss in a later context.*

Theorem 2. ***(Connection with Hamilton's Equation)*** *Let $x(s)$ be the optimal solution to the variational problem and $p(s)$ be its generalized momentum defined as $p(s) = D_vL(x'(s), x(s))$ above, then those two quantities satisfy Hamilton's equations*

$$\begin{cases} x'(s) = D_pH(p(s), x(s)) \\ p'(s) = -D_xH(p(s), x(s)) \end{cases} \tag{41}$$

for $s \in [0, t]$ and the mapping

$$s \to H(p(s), x(s)) \tag{42}$$

is constant.

Proof. Here is where the assumption that $p = D_vL(v, x)$ has unique smooth solution $v = v(p, x)$ comes in. By such assumption, we conclude that $v(p(s), x(s)) = x'(s)$.

After noticing this fact, we are left with pure calculations for D_pH, D_xH. By definition, $H(p, x) = p \cdot v(p, x) - L(v(p, x), x)$, so

$$H_{p_i}(p, x) = \sum_{j=1, j\neq i}^{n} p_j \cdot v^j_{p_i}(p, x) + v^i(p, x) + p_i \cdot v^i_{p_i}(p, x) - D_vL(v(p, x), x) \cdot D_{p_i}v(p, x) \tag{43}$$

$$= \sum_{j=1}^{n} [p_j \cdot v^j_{p_i}(p, x) - L_{v_j}(v(p, x), x) \cdot v^j_{p_i}(p, x)] + v^i(p, x) \tag{44}$$

$$= \sum_{j=1}^{n} [p_j - L_{v_j}(v(p, x), x)] \cdot v^j_{p_i}(p, x) + v^i(p, x) \tag{45}$$

$$= v^i(p, x) \tag{46}$$

since $p = D_vL(v, x)$ by the definition of v. As a result,

$$H_{p_i}(p(s), x(s)) = v^i(p(s), x(s)) = [x_i(s)]' \tag{47}$$

proved how the first n equations come.

For the next n equations, the calculation is similar

$$H_{x_i}(p,x) = \sum_{j=1}^{n} p_j v^j_{x_i}(p,x) - D_v L(v(p,x),x) \cdot D_{x_i} v(p,x) - D_x L(v(p,x),x) \tag{48}$$

$$= \sum_{j=1}^{n} [p_j v^j_{x_i}(p,x) - p_j \cdot v^j_{x_i}(p,x)] - D_x L(v(p,x),x) \tag{49}$$

$$= -D_x L(v(p,x),x) \tag{50}$$

by applying the definition of v once more. As a result,

$$H_{x_i}(p(s),x(s)) = -D_x L(v(p(s),x(s)),x(s)) = -D_x L(x'(s),x(s)) \tag{51}$$

proves the Hamilton's equations.

Moreover,

$$\frac{d}{ds} H(p(s),x(s)) = D_p H(p(s),x(s)) \cdot p'(s) + D_x H(p(s),x(s)) \cdot x'(s) \tag{52}$$

$$= x'(s) \cdot p'(s) - p'(s) \cdot x'(s) \tag{53}$$

$$= 0 \tag{54}$$

and this is telling us that the Hamiltonian is invariant w.r.t. time. □

Remark. *To briefly conclude what we have talked about in this section, we start from introducing Lagrangian as the "running loss function" of the variational problem and hope to find the optimal path $x(s)$ minimizing the loss. Such optimal path shall then satisfy the Euler-Lagrange equation consisting of n second-order ODEs.*

From the Euler-Lagrange equations, one can further introduce the generalized momentum $p(s)$ and the velocity $v(p,x)$ as the unique smooth solution to $p = D_v L(v,x)$ (such $v(s) = x'(s)$ is the unique velocity such that the generalized momentum is the given p). The Hamiltonian is defined and the optimal path $x(s)$ and the generalized momentum $p(s)$ must satisfy the Hamilton's equation. Moreover, ***the Hamiltonian won't change as time goes by on the optimal path****.*

As a result, we have interpreted the meaning of the Hamilton equations derived from the method of characteristics.

Legendre Transform & Frenchel Conjugate

Now let's turn back to HJE

$$\begin{cases} u_t + H(Du) = 0 & \text{in } \mathbb{R}^n \times (0,\infty) \\ u = g & \text{on } \mathbb{R}^n \times \{t=0\} \end{cases} \tag{55}$$

with **the dependence on x of Hamiltonian H cancelled**. Now the Lagrangian $L(v)$ only depends on v. Let's **assume that the Lagrangian is a convex function with** $\lim_{||v||\to\infty} \frac{L(v)}{||v||} = +\infty$ so of course it's continuous.

The **Legendre transform** provides the **Frenchel conjugate** of the Lagrangian as

$$L^*(p) = \sup_{v\in\mathbb{R}^n} \{p\cdot v - L(v)\} \tag{56}$$

The motivation of considering Frenchel conjugate comes from the fact that in previous discussions the Hamiltonian is defined as $H(p,x) = p\cdot v(p,x) - L(v(p,x),x)$, a form very similar to the conjugate of Lagrangian. To figure out the relationship between Hamiltonian and Lagrangian, notice that under the assumptions for Lagrangian, $p\cdot v - L(v)$ is concave and continuous in v. For each fixed $p\in\mathbb{R}^n$,

$$\frac{p\cdot v - L(v)}{||v||} = p\cdot\frac{v}{||v||} - \frac{L(v)}{||v||} \to -\infty \ (||v||\to\infty) \tag{57}$$

so there must **exist** $v^*\in\mathbb{R}^n$ **such that the sup can be attained**, i.e. $L^*(p) = p\cdot v^* - L(v^*)$. Note that **if L is differentiable at** v^*, then

$$p - DL(v^*) = 0 \tag{58}$$

since v^* achieves the sup. This gives us the equation $p = DL(v^*)$ which is just the definition equation for velocity $v(p)$ in the context above. As a result, $v(p) = v^*$ is the solution (although no uniqueness ensured). Replacing v^* with the velocity $v(p)$ one can see

$$p\cdot v(p) - L(v(p)) = L^*(p) \tag{59}$$

and the LHS is an analogue to the definition of the Hamiltonian at p! Heuristically, this gives rise to the convex duality construction of Lagrangian and Hamiltonian.

Theorem 3. ***(Convex Duality of Lagrangian and Hamiltonian)*** *Assume that Lagrangian $L = L(v)$ is convex and $\lim_{||v||\to\infty}\frac{L(v)}{||v||} = +\infty$ and* ***define*** *Hamiltonian $H = L^*$, then H is still convex, $\lim_{||p||\to\infty}\frac{H(p)}{||p||} = +\infty$ and $L = H^*$.*

In particular, when H is differentiable at p and L is differentiable at v, then the followings are equivalent:

$$\begin{cases} p \cdot v = L(v) + H(p) \\ p = DL(v) \\ v = DH(p) \end{cases} \tag{60}$$

Proof. Note that $H = L^*$ so $H^* = L^{**}$. Note that since L is a convex and closed function, its Frenchel conjugate must be itself (since double Frenchel conjugate gives the convex envelope), so $H^* = L$ is still convex and closed.

Notice that $H(p) = \sup_{v\in\mathbb{R}^n} \{p \cdot v - L(v)\}$, so

$$\forall \lambda > 0, H(p) \geq p \cdot \lambda \frac{p}{||p||} - L\left(\lambda \frac{p}{||p||}\right) \tag{61}$$

$$\geq \lambda ||p|| - \sup_{B(0,\lambda)} L \tag{62}$$

it's then obvious that $\lim_{||p||\to\infty} \frac{H(p)}{||p||} \geq \lambda$, so $\lim_{||p||\to\infty} \frac{H(p)}{||p||} = +\infty$.

When H is differentiable at p and L is differentiable at v, note that if $p \cdot v = L(v) + H(p)$ then v is achieving the sup in $H(p) = \sup_{v\in\mathbb{R}^n} \{p \cdot v - L(v)\}$ so

$$p - DL(v) = 0 \tag{63}$$

and p is achieving the sup in $L(v) = \sup_{p\in\mathbb{R}^n} \{p \cdot v - H(p)\}$ so

$$v - DH(p) = 0 \tag{64}$$

Conversely, if $p = DL(v)$, then it's true that $H(p) = p \cdot v - L(v)$ so it's proved. □

Remark. *Consider the previous example that*

$$L(v) = \frac{1}{2} m ||v||^2 \tag{65}$$

then $H(p) = \sup_{v\in\mathbb{R}^n} \left\{p \cdot v - \frac{1}{2} m ||v||^2\right\}$ and the sup is achieved at $v^* = \frac{1}{m} p$

$$H(p) = \frac{1}{2m} ||p||^2 \tag{66}$$

if $p = DL(v) = mv$, then the Hamiltonian is actually

$$H(p) = \frac{1}{2} m ||v||^2 \tag{67}$$

which is equal to the Lagrangian when there's no potential and $H(p) + L(v) = p \cdot v$.

Remark. *Let's compute some more examples to illustrate the connection between Lagrangian and Hamiltonian.*

Consider $H(p) = \frac{1}{r}||p||^r$ $(1 < r < \infty)$, *so*

$$L(v) = \sup_{p\in\mathbb{R}^n} \{p \cdot v - H(p)\} \tag{68}$$

and the sup is achieved when $v = p \cdot ||p||^{r-2}$, *so* v *is parallel to* p *with* $p = kv$ $(k > 0)$. *Plug in to find*

$$L(v) = \sup_{k>0} \left\{k||v||^2 - \frac{k^r}{r}||v||^r\right\} \tag{69}$$

and take another derivative w.r.t. k *to find that the sup is achieved when* $k = ||v||^{\frac{2-r}{r-1}}$, *so*

$$L(v) = \frac{r-1}{r}||v||^{\frac{r}{r-1}} \tag{70}$$

$$= \frac{1}{s}||v||^s \tag{71}$$

where $\frac{1}{s} + \frac{1}{r} = 1$, *so* s *is the Holder conjugate of* r.

Consider $H(p) = \frac{1}{2}p^T Ap + b \cdot p$, *where* A *is symmetric, positive definite and* $b \in \mathbb{R}^n$, *then*

$$L(v) = \sup_{p\in\mathbb{R}^n} \{p \cdot v - H(p)\} \tag{72}$$

and the sup is achieved when $p = A^{-1}(v - b)$, *so*

$$L(v) = \frac{1}{2}(v - b)^T A^{-1}(v - b) \tag{73}$$

Remark. *For convex function, one can define the subdifferential of* H *at* p *so that the Frenchel inequality holds*

$$H(p) + L(v) \geq p \cdot v \tag{74}$$

and the equality holds if and only if $v \in \partial H(p)$ *if and only if* $p \in \partial L(v)$, *a generalization of the conclusion in the theorem above.*

Hopf-Lax Formula

We still consider the HJE with Hamiltonian H not depend on x but only depends on Du. The characteristic ODEs then become

$$\begin{cases} p'(s) = 0 \\ z'(s) = DH(p(s)) \cdot p(s) - H(p(s)) \\ x'(s) = DH(p(s)) \end{cases} \tag{75}$$

with the equation for $p'(s), x'(s)$ being Hamilton's equations. Note that since $x'(s) = DH(p(s))$, by the theorem we have proved above, $L(x'(s)) + H(p(s)) = p \cdot x'(s)$. So **the equation of $z'(s)$ is describing the fact that** $z'(s) = L(x'(s))$. From the method of characteristics,

$$z(t) = u(x(t), t) = \int_0^t L(x'(s))\, ds + g(x(0)) \tag{76}$$

since $z(0) = u(x(0), 0) = g(x(0))$ by the initial value condition, providing us an ansatz of the solution. However, this construction of the solution $u(x,t)$ assumes the smoothness of the solution, which is often not the case for HJE. To think about modifying the construction of the solution such that it also works for non-smooth solution $u(x,t)$, we notice that

$$\int_0^t L(x'(s))\, ds \tag{77}$$

is the "running loss function" of the variational problem we have mentioned above and $x(s)$ is the optimal path found in that problem. As a result, we can think about **defining**

$$u(x,t) \overset{def}{=} \inf_w \left\{ \int_0^t L(w'(s))\, ds + g(w(0)) : w : [0,t] \to \mathbb{R}^n, w \in C^1, w(t) = x \right\} \tag{78}$$

as the optimal "loss" determined by the Lagrangian among all paths that hits x at time t. To see how this works as the solution to the HJE, refer to the following theorem. We **assume that H is smooth and convex with $\lim_{||p|| \to \infty} \frac{H(p)}{||p||} = +\infty$ and $g : \mathbb{R}^n \to \mathbb{R}$ is Lipschitz in the following context**.

Theorem 4. ***(Hopf-Lax Formula)*** *For fixed $x \in \mathbb{R}^n, t > 0$,*

$$u(x,t) = \inf_w \left\{ \int_0^t L(w'(s))\, ds + g(w(0)) : w : [0,t] \to \mathbb{R}^n, w \in C^1, w(t) = x \right\} \tag{79}$$

$$= \inf_{y \in \mathbb{R}^n} \left\{ tL\left(\frac{x-y}{t}\right) + g(y) \right\} \tag{80}$$

Proof. Consider $\forall y \in \mathbb{R}^n$ and the path $w(s) = y + \frac{s}{t}(x-y)$ so $w(t) = x$ (constructed based on $\frac{x-y}{t}$ inside the Lagrangian), it's obvious that

$$u(x,t) \le \int_0^t L\left(\frac{x-y}{t}\right) ds + g(y) \tag{81}$$

so by taking inf w.r.t. y on both sides

$$u(x,t) \le \inf_{y \in \mathbb{R}^n} \left\{ tL\left(\frac{x-y}{t}\right) + g(y) \right\} \tag{82}$$

Conversely, for any C^1 path w such that $w(t) = x$, take $y = w(0)$

$$tL\left(\frac{x-y}{t}\right) + g(y) = tL\left(\frac{x-w(0)}{t}\right) + g(w(0)) \tag{83}$$

$$= tL\left(\frac{1}{t}\int_0^t w(s)\,ds\right) + g(w(0)) \tag{84}$$

$$\leq \int_0^t L(w'(s))\,ds + g(w(0)) \tag{85}$$

because of Jensen's inequality applied for $\frac{1}{t}\int_0^t f(s)\,ds$, the integral average of f on $[0,t]$

$$\frac{1}{t}\int_0^t L\left(w'(s)\right)\,ds \geq L\left(\frac{1}{t}\int_0^t w'(s)\,ds\right) \tag{86}$$

by taking inf w.r.t. all paths w on both sides, one can conclude that

$$u(x,t) \geq \inf_{y\in\mathbb{R}^n}\left\{tL\left(\frac{x-y}{t}\right) + g(y)\right\} \tag{87}$$

so the theorem is proved. □

Remark. *The shifting from the ansatz $u(x(t),t) = \int_0^t L(x'(s))\,ds + g(x(0))$ to Hopf-Lax formula is critical! The main thought comes from the variational problem viewing $x(s)$ as the optimal path and the integral of Lagrangian as running loss.*

Actually, from another perspective, one may be able to see the spirit of stochastic control out of the Hopf-Lax formula. Notice that w can be view as a stochastic process instead of a deterministic function, and the $\int_0^t L(w'(s))\,ds$ can be viewed as a running loss with $g(w(0))$ as terminal loss (conditional on the filtration $\mathscr{F}_t$, i.e. all information available until time t, that's why the domain of w is $[0,t]$. Then $u(x,t)$ is essentially a value function conditioning on $w(t) = x$, i.e. the process passes through x at time t. In such sense, ***HJE is actually characterizing the value function of a stochastic control problem in a deterministic way!***

Remark. ***In Hopf-Lax formula, the inf can always be attained****. Note that $f(y) = tL\left(\frac{x-y}{t}\right) + g(y)$ is continuous in y and*

$$\frac{f(y)}{||y||} = \frac{L\left(\frac{x-y}{t}\right)}{\frac{||y||}{t}} + \frac{g(y)}{||y||} \tag{88}$$

with $L = H^*$ so $\lim_{||v||\to\infty}\frac{L(v)}{||v||} = +\infty$ and since g is Lipschitz, $\frac{g(y)}{||y||} \leq \frac{g(0)+Lips(g)||y||}{||y||} \leq Lips(g) + \varepsilon$ for large enough $||y||$. As a result,

$$\frac{f(y)}{||y||} \to +\infty\ (||y|| \to \infty) \tag{89}$$

combining with continuity, we see that the minimum of $f(y)$ must be attained by some $y \in \mathbb{R}^n$.

Hopf-Lax Formula as Solution to HJE

Now let's argue that the heuristic definition of such $u(x,t)$ by the Hopf-Lax formula is actually a solution to HJE. In order to prove this, let's first consider some useful propositions.

Theorem 5. ***(Flow Property)*** *For each $x \in \mathbb{R}^n$ and $s \in [0,t]$,*

$$u(x,t) = \inf_{y\in\mathbb{R}^n} \left\{(t-s)L\left(\frac{x-y}{t-s}\right) + u(y,s)\right\} \tag{90}$$

Proof. Let's start by noticing that for $\forall y \in \mathbb{R}^n, s \in [0,t]$, there always exists $z \in \mathbb{R}^n$ such that the inf in Hopf-Lax formula is attained, i.e.

$$u(y,s) = sL\left(\frac{y-z}{s}\right) + g(z) \tag{91}$$

in order to connect it with $\frac{x-y}{t-s}$, consider the convex representation and apply the convexity of L that

$$\frac{t-s}{t}\frac{x-y}{t-s} + \frac{s}{t}\frac{y-z}{s} = \frac{x-z}{t} \tag{92}$$

$$\frac{t-s}{t}L\left(\frac{x-y}{t-s}\right) + \frac{s}{t}L\left(\frac{y-z}{s}\right) \geq L\left(\frac{x-z}{t}\right) \tag{93}$$

so that

$$(t-s)L\left(\frac{x-y}{t-s}\right) + u(y,s) = (t-s)L\left(\frac{x-y}{t-s}\right) + sL\left(\frac{y-z}{s}\right) + g(z) \geq tL\left(\frac{x-z}{t}\right) + g(z) \tag{94}$$

take inf w.r.t. y on both sides, one would see that

$$\inf_{y\in\mathbb{R}^n} \left\{(t-s)L\left(\frac{x-y}{t-s}\right) + u(y,s)\right\} \geq tL\left(\frac{x-z}{t}\right) + g(z) \geq u(x,t) \tag{95}$$

On the other hand, let's try to find $y \in \mathbb{R}^n$ such that $(t-s)L\left(\frac{x-y}{t-s}\right) + u(y,s) \leq u(x,t)$. Apply the Hopf-Lax formula again to find $w \in \mathbb{R}^n$ such that $u(x,t) = tL\left(\frac{x-w}{t}\right) + g(w)$. Consider applying the convexity of L again, to set

$$y = \frac{s}{t}x + \frac{t-s}{t}w \tag{96}$$

$$\frac{x-y}{t-s} = \frac{x-w}{t} \tag{97}$$

and apply Hopf-Lax formula for $u(y,s)$ once more to find

$$(t-s)L\left(\frac{x-y}{t-s}\right) + u(y,s) \leq (t-s)L\left(\frac{x-w}{t}\right) + u(y,s) \leq (t-s)L\left(\frac{x-w}{t}\right) + sL\left(\frac{y-w}{s}\right) + g(w) \tag{98}$$

note that $\frac{y-w}{s} = \frac{x-w}{t}$, so

$$(t-s)L\left(\frac{x-y}{t-s}\right) + u(y,s) \leq tL\left(\frac{x-w}{t}\right) + g(w) = u(x,t) \tag{99}$$

by taking inf w.r.t. y on both sides, we proved the conclusion. □

Remark. *Note that **the inf in this theorem can always be attained**. This requires proving the fact that $y \to u(y,s)$ is continuous, which will be proved in a later context.*

Remark. *The reason why we are calling this property the flow property is that this is telling us that we can act as if we are starting at time $s < t$ with initial value $u(y,s)$. Then the Hopf-Lax formula still holds for such problem and will generate the same u as what we would derive with an initial value condition at time 0. This is actually very similar to the flow property of diffusion process.*

Under the assumption that g is Lipschitz, one would see that such u is also Lipschitz in $\mathbb{R}^n \times [0,\infty)$ and it agrees with the initial value condition g, i.e. $\forall x \in \mathbb{R}^n, u(x,0) = g(x)$.

Theorem 6. ***(Lipschitz Continuity)*** *Such u is Lipschitz in $\mathbb{R}^n \times [0,\infty)$, and $\forall x \in \mathbb{R}^n, u(x,0) = g(x)$.*

Proof. First prove that $u(x,t)$ is Lipschitz in x. By Hopf-Lax formula, there exists $y \in \mathbb{R}^n$ such that $u(x,t) = tL\left(\frac{x-y}{t}\right) + g(y)$. As a result, for $\forall x, x' \in \mathbb{R}^n$,

$$u(x',t) - u(x,t) = \inf_z \left\{tL\left(\frac{x'-z}{t}\right) + g(z)\right\} - tL\left(\frac{x-y}{t}\right) - g(y) \tag{100}$$

$$\leq tL\left(\frac{x'-(x'-x+y)}{t}\right) + g(x'-x+y) - tL\left(\frac{x-y}{t}\right) - g(y) \tag{101}$$

$$= g(x'-x+y) - g(y) \leq Lips(g) \cdot ||x'-x|| \tag{102}$$

so

$$|u(x',t) - u(x,t)| \leq Lips(g) \cdot ||x'-x|| \tag{103}$$

by interchanging x and x'.

Now let's prove that u and g agree when $t = 0$. Note that by Hopf-Lax formula, $u(x,t) \leq tL(0) + g(x)$. Set $t = 0$ to find $u(x,0) \leq g(x)$. For the other direction, we would need to use the conjugacy of Lagrangian and Hamiltonian.

$$u(x,t) = \inf_{y\in\mathbb{R}^n} \left\{tL\left(\frac{x-y}{t}\right) + g(y)\right\} \tag{104}$$

$$= g(x) + \inf_{y\in\mathbb{R}^n} \left\{tL\left(\frac{x-y}{t}\right) + g(y) - g(x)\right\} \tag{105}$$

$$\geq g(x) - t\sup_{y\in\mathbb{R}^n} \left\{-L\left(\frac{x-y}{t}\right) + Lips(g) \cdot \frac{||y-x||}{t}\right\} \tag{106}$$

by setting $z = \frac{x-y}{t}$ as a new variable, one can see the structure of this sup

$$u(x,t) \geq g(x) - t \sup_{z\in\mathbb{R}^n} \left\{-L(z) + Lips(g)\cdot ||z||\right\} \tag{107}$$

in order to connect this sup with the Frenchel conjugate of Lagrangian which is the Hamiltonian, we would like to see the forms like $w\cdot z - L(z)$. That's why we view $Lips(g)\cdot ||z||$ as $Lips(g)\frac{z}{||z||}\cdot z$ with $w = Lips(g)\frac{z}{||z||}$

$$u(x,t) \geq g(x) - t \sup_{w\in B(0,Lips(g))} \sup_{z\in\mathbb{R}^n} \left\{-L(z) + w\cdot z\right\} \tag{108}$$

$$= g(x) - t \sup_{w\in B(0,Lips(g))} H(w) \tag{109}$$

and since H is continuous and convex, $\sup_{w\in B(0,Lips(g))} H(w) < \infty$, setting $t = 0$ to see

$$u(x,0) \geq g(x) \tag{110}$$

and we conclude that such u is equal to g when $t = 0$.

At last, prove that $u(x,t)$ is Lipschitz in t. For $\forall 0 < t < t'$, by the flow property,

$$u(x,t') - u(x,t) = \inf_{y\in\mathbb{R}^n} \left\{(t'-t)L\left(\frac{x-y}{t'-t}\right) + u(y,t)\right\} - u(x,t) \tag{111}$$

$$\leq (t'-t)L(0) + u(x,t) - u(x,t) \tag{112}$$

$$= (t'-t)\cdot L(0) \tag{113}$$

on the other hand, let's apply the trick above one more time

$$u(x,t') = u(x,t) + \inf_{y\in\mathbb{R}^n} \left\{(t'-t)L\left(\frac{x-y}{t'-t}\right) + u(y,t) - u(x,t)\right\} \tag{114}$$

$$\geq u(x,t) + (t'-t)\inf_{y\in\mathbb{R}^n} \left\{L\left(\frac{x-y}{t'-t}\right) - Lips(u)\cdot \frac{||y-x||}{t'-t}\right\} \tag{115}$$

consider $z = \frac{y-x}{t'-t}$ and transform $Lips(u)\cdot\frac{||y-x||}{t'-t}$ into the inner product form to see

$$u(x,t') - u(x,t) \geq -(t'-t)\sup_{z\in\mathbb{R}^n} \left\{-L(z) + Lips(u)\cdot ||z||\right\} \tag{116}$$

$$= -(t'-t) \sup_{w\in B(0,Lips(u))} \sup_{z\in\mathbb{R}^n} \left\{-L(z) + w\cdot z\right\} \tag{117}$$

$$= -(t'-t) \sup_{w\in B(0,Lips(u))} H(w) \tag{118}$$

in all, we see that

$$|u(x,t') - u(x,t)| \le C \cdot |t'-t|, C = \max\left\{|L(0)|, \sup_{w\in B(0,Lips(u))} |H(w)|\right\} \tag{119}$$

and such constant C has no dependence on x and t, that's why u is also Lipschitz w.r.t. time t.

□

Theorem 7. ***(Hopf-Lax Formula as Solution to HJE)*** *For u defined by the Hopf-Lax formula, if it's differentiable at a point (x,t), then $u_t(x,t) + H(Du(x,t)) = 0$. In particular, such u is differentiable almost everywhere and it's the solution to HJE in the almost everywhere sense.*

Proof. By Rademacher's theorem, Lipschitz function on an open subset of $\mathbb{R}^n$ is almost everywhere differentiable. So we only have to prove that HJE holds whenever u is differentiable at (x,t).

let's first calculate the directional derivative of u along any vector v. By flow property,

$$u(x+hv, t+h) = \inf_{y\in\mathbb{R}^n}\left\{hL\left(\frac{x+hv-y}{h}\right) + u(y,t)\right\} \tag{120}$$

$$\le hL(v) + u(x,t) \tag{121}$$

as a result,

$$\forall v\in\mathbb{R}^n, v\cdot Du(x,t) + u_t(x,t) = \lim_{h\to 0^+}\frac{u(x+hv,t+h)-u(x,t)}{h} \le L(v) \tag{122}$$

note that the Hamiltonian is the Frenchel conjugate of Lagrangian, so

$$u_t(x,t) + H(Du(x,t)) = u_t(x,t) + \sup_{v\in\mathbb{R}^n}\{v\cdot Du(x,t) - L(v)\} \le 0 \tag{123}$$

To prove the other side, we have to choose v in the sup carefully. By Hopf-Lax formula, there exists $z\in\mathbb{R}^n$ such that $u(x,t) = tL\left(\frac{x-z}{t}\right) + g(z)$. Take $v = \frac{x-z}{t}$ in the sup to find

$$u_t(x,t) + H(Du(x,t)) \ge u_t(x,t) + \frac{x-z}{t}\cdot Du(x,t) - L\left(\frac{x-z}{t}\right) \tag{124}$$

again we have to use finite difference to approximate the partial derivatives

$$u(x,t) - u\left(\frac{t-h}{t}x + \frac{h}{t}z, t-h\right) = tL\left(\frac{x-z}{t}\right) + g(z) - u\left(\frac{t-h}{t}x + \frac{h}{t}z, t-h\right) \tag{125}$$

$$\ge tL\left(\frac{x-z}{t}\right) + g(z) - (t-h)L\left(\frac{x-z}{t}\right) - g(z) \tag{126}$$

$$= hL\left(\frac{x-z}{t}\right) \tag{127}$$

setting $h \to 0^+$ to know

$$u_t(x,t) + \frac{x-z}{t} \cdot Du(x,t) \geq L\left(\frac{x-z}{t}\right) \tag{128}$$

Finally, we have proved that

$$u_t(x,t) + H(Du(x,t)) = 0 \tag{129}$$

□

The theorem above ends our discussion on the solution to **a particular kind of HJE (Hamiltonian only depends on Du and is convex with Lipschitz initial value condition)**. To see a direct example of the application of Hopf-Lax formula, consider the following PDE

$$\begin{cases} u_t + ||Du||^2 = 0 & in\ \mathbb{R}^n \times (0,\infty) \\ u = +\infty \cdot \mathbb{I}_{E^c} & on\ \mathbb{R}^n \times \{0\} \end{cases} \tag{130}$$

with E as a closed subset in $\mathbb{R}^n$. Now the Hamiltonian is $H(p) = ||p||^2$ so Lagrangian is its Frenchel conjugate

$$L(v) = \sup_{p\in\mathbb{R}^n} \{p \cdot v - H(p)\} = \frac{1}{4}||v||^2 \tag{131}$$

Apply the Hopf-Lax formula to find

$$u(x,t) = \inf_{y\in\mathbb{R}^n} \left\{ tL\left(\frac{x-y}{t}\right) + g(y) \right\} \tag{132}$$

$$= \inf_{y\in E} \left\{ \frac{1}{4t}||x-y||^2 \right\} \tag{133}$$

$$= \frac{1}{4t} dist^2(x,E) \tag{134}$$

the solution has something to do with the distance between x and E.

Optimal Control Problem and Hamilton-Jacobi-Bellman Equation

In this section we state the deterministic optimal control problem and find the connection between optimal control problem, HJE, Hamilton-Jacobi-Bellman equation (HJBE) and -Lax formula.

Problem Formulation

All control problems have a certain dynamics telling us how the system evolves. In optimal control problem, the dynamics is given by an ODE

$$\begin{cases} x'(s) = f(x(s), \alpha(s)) \ (s \in [t, T]) \\ x(t) = x \end{cases} \tag{135}$$

where the dynamics works in time interval $[t, T]$ with T fixed and an initial value condition given at time t. We will be varying the time t and the initial value x shortly afterwards to get a PDE describing such an optimal control problem. **Note that $x(s)$ denotes the state of the problem at time s while x denotes the initial value condition.** Viewing $x'(s)$ as $\frac{x(s+h)-x(s)}{h}$ for $h \to 0^+$, the ODE is describing how the change of state from time s to time $s + h$ happens given the current state $x(s)$ and the current **control** $\alpha(s)$. (so it's actually a **Markovian** setting since $x'(s)$ has nothing to do with $\{x(t)\}|_{t<s}$ given $x(s)$.) The control can be understood as the "action" in discrete-time Markov decision process that changes the state evolution and has something to do with the rewards.

The control is nothing complicated but a set of parameters given at each time that will change the dynamics of the system, eventually changing the state evolution of the system. Let's denote $A \subset \mathbb{R}^m$ as some given compact set consisting of all possible values the control **at a given time** $\alpha(s)$ can take. The **admissible set**

$$\mathscr{A} = \{\alpha : [0, T] \to A : \alpha(\cdot) \ measurable\} \tag{136}$$

then denotes all possible controls across the whole time interval $[0, T]$ (since control may change over time, it maps each time point to the value of control at that time point). It's then clear that the function

$$f : \mathbb{R}^n \times A \to \mathbb{R}^n \tag{137}$$

is mapping a $m + n$-dimensional vector to a n-dimensional vector. Let's **assume that f is a given bounded Lipschitz function**. This assumption is made to ensure that the ODE always has unique solution for every given control $\alpha(\cdot) \in \mathscr{A}$ denoted $x(\cdot) = x^{\alpha(\cdot)}(\cdot)$. **Our goal in optimal control problem is to find the optimal control $\alpha^*(\cdot)$ under some criteria**.

In order to define the optimality, we introduce the **cost functional** that represents the cost one has to pay selecting control α with initial value condition $x(t) = x$

$$C_{x,t}[\alpha] = \int_t^T r(x(s), \alpha(s))\, ds + g(x(T)) \tag{138}$$

here $\int_t^T r(x(s),\alpha(s))\,ds$ is the running cost and $g(x(T))$ is the terminal cost where $r:\mathbb{R}^n\times A\to\mathbb{R}, g:\mathbb{R}^n\to\mathbb{R}$ are **assumed to be bounded and Lipschitz in variable** x.

To sum up, given time $t\in[0,T]$ and the initial value condition $x(t)=x$, we want to find **the optimal control** α^* such that

$$C_{x,t}[\alpha^*]=\inf_{\alpha\in\mathscr{A}}C_{x,t}[\alpha] \tag{139}$$

Value Function

Let's consider the **value function** $u(x,t)$ as the least possible cost among all admissible control with initial value condition $x(t)=x$ (with dynamic programming approach), i.e.

$$u(x,t)=\inf_{\alpha\in\mathscr{A}}C_{x,t}[\alpha] \tag{140}$$

then we hope to find a PDE that characterizes such value function u.

Theorem 8. ***(Optimality Condition)*** *For fixed $x\in\mathbb{R}^n, 0\le t<T$ and $h>0$ such that $t+h\le T$,*

$$u(x,t)=\inf_{\alpha\in\mathscr{A}}\left\{\int_t^{t+h}r(x(s),\alpha(s))\,ds+u(x(t+h),t+h)\right\} \tag{141}$$

where $x(\cdot)=x^{\alpha(\cdot)}(\cdot)$ is the solution to the ODE for fixed control $\alpha(\cdot)$.

Proof. For any control $\alpha_1\in\mathscr{A}$, the ODE has solution $x_1(\cdot)$. Now we want to prove that LHS is less than RHS, so we have to argue that $\forall\varepsilon>0$,

$$u(x,t)\le\int_t^{t+h}r(x_1(s),\alpha_1(s))\,ds+u(x_1(t+h),t+h)+\varepsilon \tag{142}$$

In order to achieve this goal, expand the inf in the definition of $u(x,t)$ for time $t+h$ and initial value $x_1(t+h)$ to find that $\forall\varepsilon>0$, there exists $\alpha_2\in\mathscr{A}$ and the solution to the ODE for fixed control α_2 which is $x_2(\cdot)$ such that

$$u(x_1(t+h),t+h)+\varepsilon\ge C_{x_1(t+h),t+h}[\alpha_2]=\int_{t+h}^T r(x_2(s),\alpha_2(s))\,ds+g(x_2(T)) \tag{143}$$

so far, we have successfully figured out a lower bound for $u(x_1(t+h),t+h)$. To connect it with $u(x,t)$ and any control α_1, we can construct a new control α_3 that sticks to α_1 before time $t+h$ but shifts to α_2 after time $t+h$.

$$\alpha_3(s)=\alpha_1(s)\cdot\mathbb{I}_{t\le s\le t+h}+\alpha_2(s)\cdot\mathbb{I}_{t+h\le s\le T} \tag{144}$$

under our assumption, the original ODE has unique solution, and it's easy to see that

$$x_3(s)=x_1(s)\cdot\mathbb{I}_{t\le s\le t+h}+x_2(s)\cdot\mathbb{I}_{t+h\le s\le T} \tag{145}$$

is the solution to the ODE for fixed control α_3 since

$$\forall t \leq s \leq t+h, x_3'(s) = x_1'(s) = f(x_1(s), \alpha_1(s)) = f(x_3(s), \alpha_3(s)) \tag{146}$$

$$\forall t+h \leq s \leq T, x_3'(s) = x_2'(s) = f(x_2(s), \alpha_2(s)) = f(x_3(s), \alpha_3(s)) \tag{147}$$

$$x_3(t) = x_1(t) = x \tag{148}$$

now we can see that

$$u(x,t) \leq C_{x,t}[\alpha_3] \tag{149}$$

$$= \int_t^T r(x_3(s), \alpha_3(s))\, ds + g(x_3(T)) \tag{150}$$

$$= \int_t^{t+h} r(x_1(s), \alpha_1(s))\, ds + \int_{t+h}^T r(x_2(s), \alpha_2(s))\, ds + g(x_2(T)) \tag{151}$$

$$\leq \int_t^{t+h} r(x_1(s), \alpha_1(s))\, ds + u(x_1(t+h), t+h) + \varepsilon \tag{152}$$

so we have proved that

$$u(x,t) \leq \inf_{\alpha \in \mathscr{A}} \left\{ \int_t^{t+h} r(x(s), \alpha(s))\, ds + u(x(t+h), t+h) \right\} \tag{153}$$

On the other hand, $\forall \varepsilon > 0$, there exists control $\alpha_4 \in \mathscr{A}$ such that

$$u(x,t) + \varepsilon \geq C_{x,t}[\alpha_4] \tag{154}$$

$$= \int_t^T r(x_4(s), \alpha_4(s))\, ds + g(x_4(T)) \tag{155}$$

$$= \int_t^{t+h} r(x_4(s), \alpha_4(s))\, ds + \int_{t+h}^T r(x_4(s), \alpha_4(s))\, ds + g(x_4(T)) \tag{156}$$

by the inf in the definition of value function. However, by applying again the inf for $u(x_4(t+h), t+h)$

$$u(x_4(t+h), t+h) \leq C_{x_4(t+h), t+h}[\alpha_4] \tag{157}$$

$$= \int_{t+h}^T r(x_4(s), \alpha_4(s))\, ds + g(x_4(T)) \tag{158}$$

so we have proved that

$$u(x,t) + \varepsilon \geq \int_t^{t+h} r(x_4(s), \alpha_4(s))\, ds + u(x_4(t+h), t+h) \tag{159}$$

and we will find

$$u(x,t) \geq \inf_{\alpha \in \mathscr{A}} \left\{ \int_t^{t+h} r(x(s), \alpha(s))\, ds + u(x(t+h), t+h) \right\} \tag{160}$$

so the theorem is proved. $\square$

Remark. *The optimality condition is telling us a very intuitive fact: the optimal control for the process starting from x at time t, has already taken the optimal control for the process starting from $x(t+h)$ at time $t+h$ into consideration. As a result, we can view $u(x(t+h), t+h)$ as the terminal cost (only depends on the endpoint $x(t+h)$) and $\int_t^{t+h} r(x(s), \alpha(s))\, ds$ as the running cost (depends on how $x(t)$ behaves in time $[t, t+h]$). One might be able to see the Markovian structure again from this expression.*

To set up a PDE for value function $u(x,t)$, it's natural for us to prove that u is Lipschitz (so it's almost everywhere differentiable and the PDE can hold in the almost everywhere sense).

Theorem 9. ***(Boundedness and Lipschitz Continuity of Value Function)*** *The value function $u(x,t)$ under the assumptions above is bounded and Lipschitz on $\mathbb{R}^n \times [0,T]$.*

Proof. Since $u(x,t) = \inf_{\alpha \in \mathscr{A}} C_{x,t}[\alpha]$ and r, g are assumed to be bounded, it's obvious that u is also bounded

$$u(x,t) \leq \sup |r| \cdot T + \sup |g| \tag{161}$$

Now fix $t \in [0,T]$ and consider $x, \hat{x} \in \mathbb{R}$, apply the inf in the definition of value function, so $\forall \varepsilon > 0$, there exists control $\hat{\alpha}$ and $\hat{x}(s)$ as the solution to the ODE with fixed control $\hat{\alpha}$ and initial value condition $\hat{x}(t) = \hat{x}$ such that

$$u(\hat{x}, t) + \varepsilon \geq \int_t^T r(\hat{x}(s), \hat{\alpha}(s))\, ds + g(\hat{x}(T)) \tag{162}$$

so let's estimate the difference

$$u(x,t) - u(\hat{x},t) \leq u(x,t) - \int_t^T r(\hat{x}(s), \hat{\alpha}(s))\, ds - g(\hat{x}(T)) + \varepsilon \tag{163}$$

$$\leq \int_t^T r(x(s), \hat{\alpha}(s))\, ds + g(x(T)) - \int_t^T r(\hat{x}(s), \hat{\alpha}(s))\, ds - g(\hat{x}(T)) + \varepsilon \tag{164}$$

note that here we are taking $x(s)$ as the solution to the ODE with initial value condition $x(t) = x$ that

$$x'(s) = f(x(s), \hat{\alpha}(s)) \tag{165}$$

since r, g are Lipschitz with Lipschitz constant C_r, C_g,

$$u(x,t) - u(\hat{x},t) \leq C_r \int_t^T ||x(s) - \hat{x}(s)||\, ds + C_g ||x(T) - \hat{x}(T)|| + \varepsilon \tag{166}$$

in order to estimate $\int_t^T ||x(s)-\hat{x}(s)||\,ds$, note that since f is also Lipschitz with constant C_f,

$$||x'(s)-\hat{x}'(s)|| = ||f(x(s),\hat{\alpha}(s)) - f(\hat{x}(s),\hat{\alpha}(s))|| \tag{167}$$
$$\leq C_f ||x(s)-\hat{x}(s)|| \tag{168}$$

by Grownwall's inequality,

$$||x(s)-\hat{x}(s)|| \leq C||x(t)-\hat{x}(t)|| = C||x-\hat{x}|| \tag{169}$$

that's why

$$u(x,t) - u(\hat{x},t) \leq CT||x-\hat{x}|| + \varepsilon \tag{170}$$

for some constant C and thus u is Lipschitz in variable x (the other side is similar).

To prove that it's also Lipschitz in variable t, let's fix $x \in \mathbb{R}^n$ and consider $t, \hat{t} \in [0,T]$. For $\forall \varepsilon > 0$, there exists control α and the solution $x(\cdot)$ to the ODE with fixed control α such that

$$u(x,t) + \varepsilon \geq C_{x,t}[\alpha] = \int_t^T r(x(s),\alpha(s))\,ds + g(x(T)) \tag{171}$$

consider the time-shifted control $\hat{\alpha}(s) = \alpha(s+t-\hat{t})$ and $\hat{x}$ as the solution to the ODE with fixed control $\hat{\alpha}$, one may find $\hat{x}'(s) = f(\hat{x}(s),\hat{\alpha}(s))$ and $\frac{d}{ds}x(s+t-\hat{t}) = x'(s+t-\hat{t}) = f(x(s+t-\hat{t}),\alpha(s+t-\hat{t})) = f(x(s+t-\hat{t}),\hat{\alpha}(s))$. By the uniqueness of the solution, we know that $\hat{x}(s) = x(s+t-\hat{t}), \hat{x}(\hat{t}) = x(t) = x$, so

$$u(x,\hat{t}) - u(x,t) \leq u(x,\hat{t}) - \int_t^T r(x(s),\alpha(s))\,ds - g(x(T)) + \varepsilon \tag{172}$$
$$\leq \int_{\hat{t}}^T r(\hat{x}(s),\hat{\alpha}(s))\,ds + g(\hat{x}(T)) - \int_t^T r(x(s),\alpha(s))\,ds - g(x(T)) + \varepsilon \tag{173}$$
$$\leq \int_T^{T-\hat{t}+t} r(x(s),\alpha(s))\,ds + g(\hat{x}(T)) - g(x(T)) + \varepsilon \tag{174}$$
$$\leq \sup|r| \cdot |t-\hat{t}| + C_g \cdot ||\hat{x}(T) - x(T)|| + \varepsilon \tag{175}$$
$$\leq C \cdot |t-\hat{t}| + \varepsilon \tag{176}$$

since $||\hat{x}(T) - x(T)|| \leq \sup|f| \cdot |T+t-\hat{t}-T| = \sup|f| \cdot |t-\hat{t}|$ so we have proved that $u(x,t)$ is also Lipschitz in t (the other side is similar). $\square$

Hamilton-Jacobi-Bellman Equation (HJBE)

Now from the optimality condition and the Lipschitz continuity of the value function derived above, we can set up a PDE describing the evolution of value function $u(x,t)$.

Theorem 10. ***(HJBE for Value Function)*** *The value function under assumptions above satisfies the HJBE*

$$\begin{cases} u_t + \inf_{\alpha\in\mathscr{A}} \{f(x,\alpha)\cdot Du + r(x,\alpha)\} = 0 & \text{in } \mathbb{R}^n \times (0,T) \\ u = g & \text{on } \mathbb{R}^n \times \{t\} \end{cases} \tag{177}$$

Proof. When $t = T$, $u = \inf_{\alpha\in\mathscr{A}} C_{x,T}[\alpha] = \int_T^T r(x(s),\alpha(s))\,ds + g(x(T)) = g(x)$ gives the terminal condition.

When $0 < t < T$, recall the optimality condition that for $h > 0$ such that $t + h \leq T$,

$$u(x,t) = \inf_{\alpha\in\mathscr{A}} \left\{ \int_t^{t+h} r(x(s),\alpha(s))\,ds + u(x(t+h),t+h) \right\} \tag{178}$$

where $x(\cdot)$ is the solution to the ODE for fixed control α. Let's modify both sides of this property to get HJBE, be careful with the difference between x and $x(\cdot)$ since the previous one denotes the initial value while the latter one denotes the solution to the PDE

$$\frac{u(x,t) - u(x,t+h)}{h} = \inf_{\alpha\in\mathscr{A}} \left\{ \frac{1}{h}\int_t^{t+h} r(x(s),\alpha(s))\,ds + \frac{u(x(t+h),t+h) - u(x,t+h)}{h} \right\} \tag{179}$$

$$= \inf_{\alpha\in\mathscr{A}} \left\{ \frac{1}{h}\int_t^{t+h} r(x(s),\alpha(s))\,ds + \frac{u(x(t+h),t+h) - u(x(t),t+h)}{h} \right\} \tag{180}$$

setting $h \to 0^+$ on both sides to find

$$-u_t(x,t) = \inf_{\alpha\in\mathscr{A}} \{r(x(t),\alpha(t)) + Du(x(t),t)\cdot x'(t)\} \tag{181}$$

$$= \inf_{\alpha\in\mathscr{A}} \{r(x(t),\alpha(t)) + Du(x,t)\cdot f(x(t),\alpha(t))\} \tag{182}$$

now let's neglect the initial time t and initial value x to denote the PDE as

$$u_t + \inf_{\alpha\in\mathscr{A}} \{r(x,\alpha) + Du\cdot f(x,\alpha)\} = 0 \tag{183}$$

note that u being Lipschitz guarantees that the partial derivatives w.r.t. each variable exists almost everywhere. $\square$

Remark. *We can find the connection between HJE and HJBE that if we set the Hamiltonian as*

$$H(p,x) = \inf_{\alpha\in\mathscr{A}} \{f(x,\alpha)\cdot p + r(x,\alpha)\} \tag{184}$$

*then HJBE is just HJE $u_t + H(Du,x) = 0$ but with a **terminal value condition** instead of an initial value condition.*

Remark. *One may still recall the Hopf-Lax formula mentioned above to solve HJE $u_t + H(Du) = 0$ with initial value condition $u(x, 0) = g(x)$ that*

$$u(x,t) = \inf_{y\in\mathbb{R}^n} \left\{ tL\left(\frac{x-y}{t}\right) + g(y) \right\} \tag{185}$$

with the Lagrangian L as the Frenchel conjugate of the Hamiltonian H. We can verify that such $u(x, t)$ also provides us with the solution to a special kind of HJBE.

Now that HJE has initial value condition but HJBE has terminal value condition, the most natural way is to do the time reflection $v(x, t) = u(x, T - t)$ such that the terminal value condition of u actually gives the initial value condition of v. It's easy to see that

$$v(x, 0) = u(x, T) = g(x) \tag{186}$$

then notice that $v_t = -u_t, Dv = Du$, so the HJBE for u can be reformulated as the following HJE for v that

$$\begin{cases} v_t + H(Dv, x) = 0 & \text{in } \mathbb{R}^n \times (0, T) \\ v = g & \text{on } \mathbb{R}^n \times \{0\} \end{cases} \tag{187}$$

with Hamiltonian

$$H(p, x) = -\inf_{\alpha\in\mathscr{A}} \{f(x, \alpha) \cdot p + r(x, \alpha)\} \tag{188}$$

However, in order to let the Hopf-Lax formula work, we have to ***assume that*** $r(x, \alpha) = r(\alpha), f(x, \alpha) = f(\alpha)$***, i.e. both running reward and the dynamics does not depend on the state*** x*. So the HJE and the Hamiltonian becomes*

$$\begin{cases} v_t + H(Dv) = 0 & \text{in } \mathbb{R}^n \times (0, T) \\ v = g & \text{on } \mathbb{R}^n \times \{0\} \end{cases} \tag{189}$$

and

$$H(p) = -\inf_{\alpha\in\mathscr{A}} \{f(\alpha) \cdot p + r(\alpha)\} \tag{190}$$

So the Frenchel conjugate is

$$L(v) = \sup_{p\in\mathbb{R}^n} \left\{ p \cdot v + \inf_{\alpha\in\mathscr{A}} \{f(\alpha) \cdot p + r(\alpha)\} \right\} \tag{191}$$

and the solution to HJE is given by

$$v(x,t) = \inf_{y\in\mathbb{R}^n}\left\{tL\left(\frac{x-y}{t}\right)+g(y)\right\} \tag{192}$$

$$= \inf_{y\in\mathbb{R}^n}\left\{\sup_{p\in\mathbb{R}^n}\left\{p\cdot(x-y)+t\inf_{\alpha\in\mathscr{A}}\{f(\alpha)\cdot p+r(\alpha)\}\right\}+g(y)\right\} \tag{193}$$

as a result, ***the solution to HJBE*** *is*

$$u(x,t) = v(x,T-t) = \inf_{y\in\mathbb{R}^n}\left\{\sup_{p\in\mathbb{R}^n}\left\{p\cdot(x-y)+(T-t)\inf_{\alpha\in\mathscr{A}}\{f(\alpha)\cdot p+r(\alpha)\}\right\}+g(y)\right\} \tag{194}$$

under the assumption that g ***is Lipschitz,*** H ***is convex and*** $\lim_{||p||\to\infty}\frac{H(p)}{||p||}=+\infty$.

However, one might realize that although we have got an analytic solution for HJBE, the assumption that the running reward and the dynamics both do not depend on state is too strong that most of the interesting examples would not satisfy such assumption. This assumption only works well for a problem setting with a single state and many actions to be chosen, i.e. the continuous-time bandit problem but fails for most reinforcement learning problems.

Although one would not be able to solve the HJBE analytically in all cases, our previous discussion about general HJE $u_t+H(Du,x)=0$ *still provides some insights. One can consider the Hamilton's equation and the Euler-Lagrange equations associated with the HJBE.*

Infinite-Horizon Problem

Among our discussion, we are assuming that there exists some upper time limit $T<\infty$ and the dynamics works in time interval $[0,T]$. However, one can also consider the infinite-horizon problem by taking $T=\infty$. Let's adopt all same assumptions for A,f,r,g above, and consider the admissible set

$$\mathscr{A} = \{\alpha:[0,\infty)\to A : \alpha(\cdot)\ measurable\} \tag{195}$$

with $x(\cdot)$ as the unique solution to ODE

$$\begin{cases} x'(s) = f(x(s),\alpha(s)) \\ x(0)=x \end{cases} \tag{196}$$

for fixed control α. In order to ensure that the cost is well-defined on infinite time horizon, let's introduce $\lambda>0$ as continuous-time discount rate and define the cost as

$$C_x[\alpha] = \int_0^\infty e^{-\lambda s}r(x(s),\alpha(s))\,ds \tag{197}$$

and the value function as

$$u(x) = \inf_{\alpha \in \mathscr{A}} C_x[\alpha] \tag{198}$$

note that **the biggest difference is that infinite time horizon problem under the Markovian setting has time-homogeneous value function**.

Remark. *To see this, let's assume that former definition still applies*

$$C_{x,t}[\alpha] = \int_t^\infty e^{-\lambda s} r(x(s), \alpha(s))\, ds \tag{199}$$

and the value function is

$$u(x,t) = \inf_{\alpha \in \mathscr{A}} C_{x,t}[\alpha] \tag{200}$$

with the ODE having initial value condition $x(t) = x$. Now consider $\forall t > 0$,

$$C_{x,t}[\alpha] = \int_t^\infty e^{-\lambda s} r(x(s), \alpha(s))\, ds \tag{201}$$

$$= \int_0^\infty e^{-\lambda(s+t)} r(x(s+t), \alpha(s+t))\, ds \tag{202}$$

where $x'(s) = f(x(s), \alpha(s)), x(t) = x$. However, let's consider another solution $\hat{x}(s)$ to the ODE with fixed control $\hat{\alpha}(s) = \alpha(s+t)$ such that $\hat{x}'(s) = f(\hat{x}(s), \hat{\alpha}(s)), \hat{x}(0) = x$, according to the uniqueness of the solution to the ODE, we immediately know that $\hat{x}(s) = x(s+t)$. So now

$$C_{x,t}[\alpha] = \int_0^\infty e^{-\lambda(s+t)} r(x(s+t), \alpha(s+t))\, ds \tag{203}$$

$$= e^{-\lambda t} \cdot \int_0^\infty e^{-\lambda s} r(\hat{x}(s), \hat{\alpha}(s))\, ds \tag{204}$$

$$= e^{-\lambda t} \cdot C_{x,0}[\hat{\alpha}] \tag{205}$$

and by taking inf on both sides, one would see that

$$u(x,t) = e^{-\lambda t} \cdot u(x,0) \tag{206}$$

so the time t only appears in the discount factor $e^{-\lambda t}$. That's why we only need to consider $u(x,0)$ and denote it as $u(x)$ by taking the time t as 0 by default.

Under all assumptions made above, one can see that u **is bounded and if** $\lambda > Lips(f)$ **then** u **is Lipschitz**. To argue this, one do the similar thing as done in the previous proofs. $\forall x, \hat{x} \in \mathbb{R}^n, \forall \varepsilon > 0$, there exists control

$\hat{\alpha} \in \mathscr{A}$ and the solution $\hat{x}(s)$ to the ODE with fixed control $\hat{\alpha}$ and initial value condition $\hat{x}(0) = \hat{x}$ such that

$$u(\hat{x}) + \varepsilon \geq \int_0^\infty e^{-\lambda s} r(\hat{x}(s), \hat{\alpha}(s))\, ds \tag{207}$$

now by definition,

$$u(x) - u(\hat{x}) \leq u(x) - \int_0^\infty e^{-\lambda s} r(\hat{x}(s), \hat{\alpha}(s))\, ds + \varepsilon \tag{208}$$

$$\leq \int_0^\infty e^{-\lambda s} r(x(s), \hat{\alpha}(s))\, ds - \int_0^\infty e^{-\lambda s} r(\hat{x}(s), \hat{\alpha}(s))\, ds + \varepsilon \tag{209}$$

where $x(s)$ is the solution to the ODE with fixed control $\hat{\alpha}$ and initial value condition $x(0) = x$. So we know that

$$u(x) - u(\hat{x}) \leq \int_0^\infty e^{-\lambda s} [r(x(s), \hat{\alpha}(s)) - r(\hat{x}(s), \hat{\alpha}(s))]\, ds + \varepsilon \tag{210}$$

$$\leq C_r \cdot \int_0^\infty e^{-\lambda s} \cdot ||x(s) - \hat{x}(s)||\, ds + \varepsilon \tag{211}$$

and $||x'(s) - \hat{x}'(s)|| = ||f(x(s), \hat{\alpha}(s)) - f(\hat{x}(s), \hat{\alpha}(s))|| \leq C_f \cdot ||x(s) - \hat{x}(s)||$ so by Grownwall's inequality, we conclude that

$$||x(s) - \hat{x}(s)|| \leq e^{C_f s} \cdot ||x(0) - \hat{x}(0)|| = e^{C_f s} \cdot ||x - \hat{x}|| \tag{212}$$

so the estimates look like

$$u(x) - u(\hat{x}) \leq C_r \cdot ||x - \hat{x}|| \cdot \int_0^\infty e^{(C_f - \lambda)s}\, ds + \varepsilon \tag{213}$$

so when $C_f = Lips(f) < \lambda$, the integral converges and is a constant, that's why u is Lipschitz and is differentiable almost everywhere.

To get the HJBE for such value function $u(x)$, let's plug in

$$u(x,t) = e^{-\lambda t} \cdot u(x, 0) \tag{214}$$

into the HJBE we derived for general optimal control problem to see that

$$u_t(x,t)|_{t=0} = -\lambda \cdot u(x, 0) \tag{215}$$

so

$$-\lambda \cdot u(x, 0) + \inf_{\alpha \in \mathscr{A}} \{f(x, \alpha) \cdot Du + r(x, \alpha)\} = 0 \tag{216}$$

and we get **the HJBE for infinite-horizon optimal control problem**

$$\lambda u - \inf_{\alpha\in\mathscr{A}} \{f(x,\alpha)\cdot Du + r(x,\alpha)\} = 0 \tag{217}$$

for value function $u = u(x)$.

Till now, we have finished the discussion on optimal control problems. In the following context, we will talk about stochastic control problem where the dynamics is not an ODE but an SDE. One would see that the PDE approach is still similar to what we have done here but the probabilistic approach would be very different.

For the following contents, we refer to the book *Lectures on BSDEs, Stochastic Control, and Stochastic Differential Games with Financial Applications* by Rene Carmona and the book *Continuous-time Stochastic Control and Optimization with Financial Applications* by Pham.

Stochastic Control Problem, PDE Approach

Problem Setting

In the setting of stochastic control, the state process is denoted as $\{X_t\}$, a stochastic process in $\mathbb{R}^d$, generated as the solution to a SDE for given control (action) $\{\alpha_t\}$, which is also a stochastic process. Similar to the deterministic case, let's first specify the set of all admissible controls one can choose from. Note that different from the deterministic case, here we also have to specify the **measurability** of those controls, i.e. one cannot make use of the information that can only be known in the future to determine the best control for the time being.

Let's assume that the control α_t at each fixed time t can take value in A, a subset of a Polish space. Most often, we assume that $A \subset \mathbb{R}^k$ is a compact subset and $\mathscr{A}$ denotes the set of all **admissible controls**, i.e.

$$\mathscr{A} = \{\alpha = \{\alpha_t\} : \forall t \geq 0, \alpha_t \in A\} \tag{218}$$

let's denote α as the whole stochastic process $\{\alpha_t\}$ in the following context. Sometimes, there will be uniform bounded condition added for $\alpha \in \mathscr{A}$ and sometimes we would assume that

$$\mathbb{E} \int_0^T ||\alpha_t||^2 \, dt < \infty \tag{219}$$

, i.e. $\alpha \in L^2([0,T] \times \Omega)$ is in the L^2 Hilbert space of stochastic processes on time interval $[0,T]$. However, those conditions are added as required and there's no standard formulation of the admissible set.

Let's then consider the measurability of admissible controls. Assume that we are **in the finite time horizon case and the time has upper limit** T. Then for each $t \in [0,T]$, when one wants to choose the control, one obviously cannot use all the information of $\{X_t\}_{t\in[0,T]}$ since one cannot make any current decision based on future information. Let's denote $\{\mathcal{I}_t\}$ as a filtration standing for **the information available to the controller at time** t, i.e. $\alpha_t \in \mathcal{I}_t$. There are mainly four different kinds of settings for the measurability conditions of the admissible set.

- **OL (Open Loop)** The setting where $\mathcal{I}_t = \sigma\{X_0\}$ and $\alpha_t = \alpha_t(t, X_0, \{B_s\}|_{s\in[0,t]})$.
- **CLPS (Closed Loop Perfect State)** The setting where $\mathcal{I}_t = \sigma\{X_s : s \in [0,t]\}$ and $\alpha_t = \alpha_t(t, \{X_s\}|_{s\in[0,t]})$.
- **MPS (Memoryless Perfect State)** The setting where $\mathcal{I}_t = \sigma\{X_0, X_t\}$ and $\alpha_t = \alpha_t(t, X_0, X_t)$.
- **FPS (Feedback Perfect State/Markovian)** The setting where $\mathcal{I}_t = \sigma\{X_t\}$ and $\alpha_t = \alpha_t(t, X_t)$.

In OL, the information available to the controller at time t is always the initial state and the noise so far. In MPS, the information available to the controller at time t is the initial state and the current state. In FPS, the information available to the controller at time t is only the current state but the initial state can not be observed and in CLPS, all history states are known to the controller. One might be able to find that OL is the most specific setting while CLPS is the most general setting. By mentioning Markov games, we take the FPS setting by default.

Now the dynamics of the state process is given by the SDE

$$dX_t = b(t, X_t, \alpha_t)\, dt + \sigma(t, X_t, \alpha_t)\, dB_t \tag{220}$$

where the drift and diffusion coefficient $b : [0, T] \times \mathbb{R}^d \times A \to \mathbb{R}^d, \sigma : [0, T] \times \mathbb{R}^d \times A \to \mathbb{R}^{d\times m}$. So X_t **is a process in $\mathbb{R}^d$, the BM B_t here is of m-dimension** and for each given control α_t one can solve the SDE to know X_t (the choice of action changes the state evolution). For the purpose of simplicity, we want to ensure **the existence and uniqueness of the strong solution to such SDE**. Note that both coefficients depend on the control α_t, so it's natural to make some **additional assumptions to the admissible set** that

$$\mathscr{A} = \left\{\alpha : \mathbb{E}\int_0^T ||b(t, 0, \alpha_t)||^2 + ||\sigma(t, 0, \alpha_t)||^2\, dt < \infty\right\} \tag{221}$$

now we also **assume that** $b(t, x, \alpha), \sigma(t, x, \alpha)$ **are both Lipschitz in** x so the existence and uniqueness of the strong solution can be guaranteed. Let's denote $X^{t,x,\alpha} = \{X_s^{t,x,\alpha}\}\,|_{s\in[t,T]}$ as the unique solution to the following SDE with initial value condition and given control $\alpha \in \mathscr{A}$

$$\begin{cases} dX_t = b(t, X_t, \alpha_t)\, dt + \sigma(t, X_t, \alpha_t)\, dB_t \\ X_t = x \end{cases} \tag{222}$$

The **cost functional** is defined as

$$J(\alpha) = \mathbb{E}\left[\int_0^T f(s, X_s, \alpha_s)\, ds + g(X_T)\right] \tag{223}$$

consisting of two parts, the running cost and the terminal cost, and we **assume that** $f(t, x, \alpha)$ **is Lipschitz in** x. Now the objective of stochastic control problem is to **find the optimal control** $\alpha = \alpha^*$ **such that it minimizes the cost functional** $J(\alpha)$.

Remark. *Of course, such optimal control α^* does not necessarily exist. To prove the existence, one needs to show that $\mathscr{A}$ is a convex subset and J is convex, l.s.c. with compact level sets so that the existence of the minimum is ensured. However, this does not seem to be very interesting in the scope of our discussion.*

Remark. *To mention the technique of absorbing the running cost and maintaining only the terminal cost, let's consider a new process*

$$Y_t = \int_0^t f(s, X_s, \alpha_s)\, ds \tag{224}$$

so $J(\alpha) = \mathbb{E}\left[Y_T + g(X_T)\right] = \mathbb{E}\tilde{g}(X_T, Y_T)$ *if the function* $\tilde{g}$ *is defined as*

$$\tilde{g}(x, y) = y + g(x) \tag{225}$$

As a result, under the new setting, our state process becomes $\tilde{X}_t = (X_t, Y_t)$*, and the cost functional is* $J(\alpha) = \mathbb{E}\tilde{g}(\tilde{X}_T)$ *with the same set of admissible controls. However, the cost of doing this is that: (i): the increase in the dimension of state process (ii):* Y_t *has dynamics* $dY_t = f(t, X_t, \alpha_t)\, dt$ *that has no diffusion terms.*

Example: Separable Control Problem

Consider the problem where a single firm facing regulations for pollution permits in time $[0, T]$. The firm has cumulative emissions E_t up to time t generated by the SDE

$$\begin{cases} dE_t = (b_t - \xi_t)\,dt + \sigma_t\,dB_t \\ E_0 = 0 \end{cases} \tag{226}$$

where b_t is the expected rate of emission change if there's no regulation and ξ_t is the rate of abatement chosen by the firm (so it's an action/control). However, the larger rate of abatement the form chooses, the less it can produce, so there is a cost function $c : \mathbb{R} \to \mathbb{R}$ characterizing the cost of lowering the emission. On the other hand, the form can also choose to hold θ_t quantity of pollution permits at time t, with Y_t characterizing the price of each pollution permit (there is an allowance market where firms can trade permits). At last, our goal is to figure out the best control ξ^*, θ^* such that the utility of the firm is maximized for a given utility function U.

Now we make assumptions that b_t, σ_t are adapted and bounded, c is C^1, nondecreasing, strictly convex and $c'(-\infty) = -\infty, c'(+\infty) = +\infty, c(0) = 0$, U is C^1, increasing, strictly concave and $U'(-\infty) = -\infty, U'(+\infty) = +\infty$ (the Inada condition). Note that here BM B_t and E_t are both 1-dimensional.

Let's denote X_T as the total wealth of the company at terminal time T with initial wealth $X_0 = x$, then

$$X_T = x + \int_0^T \theta_t\,dY_t - \int_0^T c(\xi_t)\,dt - E_T Y_T \tag{227}$$

here the second term on RHS stands for the wealth the firm gets in the allowance market by trading permits through time $[0, T]$, the third term on RHS is the cost in production caused by the abatement in the emission, and the last term on RHS is the final cost to eliminate all emissions with permits (there's E_T emissions altogether and each permit costs Y_T). Our goal is to find the optimal control ξ^*, θ^* such that

$$\mathbb{E}U\left(X_T^{\xi^*,\theta^*}\right) = \sup_{(\xi,\theta)\in\mathscr{A}} \mathbb{E}U\left(X_T^{\xi,\theta}\right) \tag{228}$$

the admissible set here only requires the integrability condition

$$\mathbb{E}\int_0^T ||b_t - \xi_t||^2 + ||\sigma_t||^2\,dt < \infty \tag{229}$$

Let's prove that **the optimal abatement strategy is**

$$\xi_t^* = (c')^{-1}(Y_t) \tag{230}$$

let's rewrite the terminal wealth by replacing E_t

$$E_T Y_T = Y_T\left(\int_0^T (b_t-\xi_t)\,dt + \int_0^T \sigma_t\,dB_t\right) \tag{231}$$
$$= Y_T\left(\int_0^T b_t\,dt + \int_0^T \sigma_t\,dB_t\right) - Y_T\int_0^T \xi_t\,dt \tag{232}$$

note that

$$\int_0^T (Y_T-Y_t)\xi_t\,dt = \int_0^T \left(\int_t^T dY_s\right)\xi_t\,dt \tag{233}$$
$$= \int_0^T \left(\int_0^s \xi_t\,dt\right)dY_s \tag{234}$$

plug in to find

$$X_T = x - \int_0^T c(\xi_t)\,dt + \int_0^T \theta_t\,dY_t - Y_T\left(\int_0^T b_t\,dt + \int_0^T \sigma_t\,dB_t\right) + \int_0^T Y_t\xi_t\,dt + \int_0^T\left(\int_0^s \xi_t\,dt\right)dY_s \tag{235}$$
$$= x - \int_0^T [c(\xi_t)-Y_t\xi_t]\,dt + \int_0^T \theta_t\,dY_t - Y_T\int_0^T b_t\,dt - Y_T\int_0^T \sigma_t\,dB_t + \int_0^T\left(\int_0^t \xi_s\,ds\right)dY_t \tag{236}$$
$$= x - \int_0^T [c(\xi_t)-Y_t\xi_t]\,dt + \int_0^T \left[\theta_t + \int_0^t \xi_s\,ds\right]dY_t - Y_T\int_0^T b_t\,dt - Y_T\int_0^T \sigma_t\,dB_t \tag{237}$$

call the first two terms on RHS as B_T^ξ and the remaining terms on RHS as $A_T^{\tilde\theta}$ with the new control defined as $\tilde\theta_t = \theta_t + \int_0^t \xi_s\,ds$, so now

$$\begin{cases} B_T^\xi = x - \int_0^T [c(\xi_t)-Y_t\xi_t]\,dt \\ A_T^{\tilde\theta} = \int_0^T \tilde\theta_t\,dY_t - Y_T\int_0^T b_t\,dt - Y_T\int_0^T \sigma_t\,dB_t \end{cases} \tag{238}$$

those two parts are separated such that B_T^ξ has nothing to do with $\tilde\theta$ and $A_T^{\tilde\theta}$ has nothing to do with ξ if we see $\xi,\tilde\theta$ as two independent controls (although they are actually not since the definition of $\tilde\theta$ contains ξ). However, we can notice that when (θ,ξ) traverses through the admissible set $\mathscr{A}$, $(\tilde\theta,\xi)$ also traverses through the admissible set $\mathscr{A}$ and vice versa. As a result,

$$\sup_{(\xi,\theta)\in\mathscr{A}} \mathbb{E}U\left(X_T^{\xi,\theta}\right) = \sup_{(\xi,\tilde\theta)\in\mathscr{A}} \mathbb{E}U\left(X_T^{\xi,\theta}\right) \tag{239}$$
$$= \sup_{(\xi,\tilde\theta)\in\mathscr{A}} \mathbb{E}U\left(A_T^{\tilde\theta} + B_T^\xi\right) \tag{240}$$
$$= \sup_{\tilde\theta\in\mathscr{A}}\sup_{\xi\in\mathscr{A}} \mathbb{E}U\left(A_T^{\tilde\theta} + B_T^\xi\right) \tag{241}$$

so the optimal abatement rate ξ_t^* is the ξ_t that maximizes B_T^ξ (under the maximization of xi, $A_T^{\tilde{\theta}}$ is a constant and note that the utility is increasing, under the assumptions, the maximum exists and $(c')^{-1}$ exists).

$$\xi_t^* = \arg\max_{\xi_t} \left\{ x - \int_0^T [c(\xi_t) - Y_t\xi_t]\, dt \right\} \tag{242}$$

$$\xi_t^* = (c')^{-1}(Y_t) \in \mathcal{I}_t \tag{243}$$

Remark. *The trick applied in this example is to set up a new control such that the wealth is a* ***separable*** *and argue that the old set of controls traverse through the admissible set if and only if the new set of controls traverse through the admissible set. As a result, the new controls can be seen as independent controls and two maximization can be dealt with separately.*

Note that one has to verify that the optimal control one get satisfies measurability requirements. For example, in this example, we are taking the Markovian setting so ξ_t^* can only depend on the value of all observable processes at time t.

Remark. *To get the intuition of such optimal abatement rate, it's telling us that on observing the price of the pollution permit Y_t at time t, the firm shall always make sure that* ***the marginal production cost*** *$c'(\xi_t)$* ***is equal to the marginal emission cost*** *Y_t. In economics, it's rational to only compare the marginal so we would get the same conclusion from intuition.*

Example: Separable Control Problem

Let's use a slightly different example to illustrate the same trick once again. Now still consider a single firm with regulation for emission allowances. Now the firm produce a source with price P_t following BS model such that

$$\frac{dP_t}{P_t} = \mu(P_t)\, dt + \sigma(P_t)\, dB_t \tag{244}$$

at each time the form can choose its rate of production q_t with production costs $c(q_t)$. Similar to the example above, the form has to buy permits for all the emission it produces. The price of the permit is denoted as Y_t and now the cumulative emission until time t, denoted E_t, is proportional to the production amount until time t for fixed $\varepsilon > 0$

$$E_t = \varepsilon Q_t, E_0 = 0 \tag{245}$$

$$Q_t = \int_0^t q_s\, ds \tag{246}$$

The firm has to decide θ_t, the quantity of permit to hold and q_t, the rate of production at time t, so the control is made up of the pair (θ_t, q_t). Let's still use X_T for the total wealth of the firm at time T with initial wealth $X_0 = x$, then

$$X_T = x + \int_0^T P_t q_t \, dt - \int_0^T c(q_t) \, dt + \int_0^T \theta_t \, dY_t - E_T Y_T \tag{247}$$

$$= x + \int_0^T P_t q_t \, dt - \int_0^T c(q_t) \, dt + \int_0^T \theta_t \, dY_t - \varepsilon Q_T Y_T \tag{248}$$

the utility function U is provided and we wish to find optimal control θ^*, q^* to maximize the expected terminal utility

$$\mathbb{E}U\left(X_T^{\theta^*, q^*}\right) = \sup_{\theta, q \in \mathscr{A}} \mathbb{E}U\left(X_T^{\theta, q}\right) \tag{249}$$

Likewise, we make the following assumptions that μ, σ are C^1 with bounded derivatives, cost function c is C^1 and strictly convex, satisfies the Inada condition, i.e. $c'(-\infty) = -\infty, c'(+\infty) = +\infty$, and the utility function U is C^1, increasing, strictly concave and satisfy the Inada condition, i.e. $U'(-\infty) = -\infty, U'(+\infty) = +\infty$. The admissible set of controls only have adaptability and integrability conditions as stated in the previous context.

Now **the optimal production strategy** should be

$$q_t^* = (c')^{-1}(P_t - \varepsilon Y_t) \tag{250}$$

since the marginal cost of producing is $c'(q_t) + \varepsilon Y_t$ (the rising in cost and the need to buy permit for increased emission) and the marginal profit of producing is P_t. By previous explanations on the intuitions, it's easy to see that optimal control is achieved when **the marginals are equal**.

Let's apply the same trick of separating two control variables here by transforming the term $Q_T Y_T$ using Ito formula (note that Q_t has finite variation)

$$d(Q_t Y_t) = Q_t \, dY_t + Y_t \, dQ_t \tag{251}$$

$$Q_T Y_T = \int_0^T Q_t \, dY_t + \int_0^T Y_t \, dQ_t \tag{252}$$

$$= \int_0^T \left(\int_0^t q_s \, ds\right) dY_t + \int_0^T Y_t q_t \, dt \tag{253}$$

plug into the expression for X_T to see that

$$X_T = x + \int_0^T [(P_t - \varepsilon Y_t) q_t - c(q_t)] \, dt + \int_0^T \theta_t \, dY_t - \varepsilon \int_0^T \left(\int_0^t q_s \, ds\right) dY_t \tag{254}$$

$$= x + \int_0^T [(P_t - \varepsilon Y_t) q_t - c(q_t)] \, dt + \int_0^T \left[\theta_t - \varepsilon \int_0^t q_s \, ds\right] dY_t \tag{255}$$

denote the new control $\tilde{\theta}_t = \theta_t - \varepsilon \int_0^t q_s \, ds$ to find that

$$X_T = x + \int_0^T [(P_t - \varepsilon Y_t)q_t - c(q_t)] \, dt + \int_0^T \tilde{\theta}_t \, dY_t \tag{256}$$

and separate it into two parts

$$\begin{cases} A_T^{\tilde{\theta}} = \int_0^T \tilde{\theta}_t \, dY_t \\ B_T^q = x + \int_0^T [(P_t - \varepsilon Y_t)q_t - c(q_t)] \, dt \end{cases} \tag{257}$$

note that when (θ, q) traverse through the admissible set, so does $(\tilde{\theta}, q)$, so

$$\begin{aligned} \sup_{\theta, q \in \mathscr{A}} \mathbb{E} U\left(X_T^{\theta,q}\right) &= \sup_{\tilde{\theta}, q \in \mathscr{A}} \mathbb{E} U\left(X_T^{\theta,q}\right) && (258) \\ &= \sup_{\tilde{\theta}, q \in \mathscr{A}} \mathbb{E} U\left(A_T^{\tilde{\theta}} + B_T^q\right) && (259) \\ &= \sup_{\tilde{\theta} \in \mathscr{A}} \sup_{q \in \mathscr{A}} \mathbb{E} U\left(A_T^{\tilde{\theta}} + B_T^q\right) && (260) \end{aligned}$$

and the optimal production strategy will be attained when B_T^q attains its maximum (since now $\tilde{\theta}, q$ are considered as independent controls and utility function is increasing, with B_T^q only depending on q and $A_T^{\tilde{\theta}}$ only depending on $\tilde{\theta}$)

$$\begin{aligned} q_t^* &= \arg\max_{q_t} \left\{ x + \int_0^T [(P_t - \varepsilon Y_t)q_t - c(q_t)] \, dt \right\} && (261) \\ q_t^* &= (c')^{-1}(P_t - \varepsilon Y_t) \in \mathcal{I}_t && (262) \end{aligned}$$

we can check that at time t, under the Markovian setting, P_t, Y_t are observable to the controller so $q_t^* \in \mathcal{I}_t$ satisfies the measurability condition.

Remark. *One might hope to use the similar technique to figure out the optimal quantity of permit to hold since*

$$\tilde{\theta}_t^* = \arg\max_{\tilde{\theta}_t} \left\{ \int_0^T \tilde{\theta}_t \, dY_t \right\} \tag{263}$$

however, one may find that by setting the derivative w.r.t. $\tilde{\theta}_t$ as 0, one cannot find the optimal $\tilde{\theta}_t$ because of the measurability issue (Y_T is not known at time t). So one cannot find an admissible control from this problem as we have done for q_t^* and one has to consider instead

$$\tilde{\theta}_t^* = \arg\max_{\tilde{\theta}_t} \left\{ \mathbb{E} U\left(\int_0^T \tilde{\theta}_t \, dY_t + B_T^{q^*} \right) \right\} \tag{264}$$

so by taking the derivative, one would get

$$\mathbb{E}\left[U'\left(\int_0^T \tilde{\theta}_t^*\, dY_t + B_T^{q^*}\right)\cdot (Y_T - Y_0)\right] = 0 \tag{265}$$

and we will see that $\int_0^T \tilde{\theta}_t^*\, dY_t$ *has something to do with the process* Y *after time* t*, so the optimal control* $\tilde{\theta}_t^*$ *is hard to figure out (especially to ensure the measurability). This is telling us that the two examples shown above have easy and intuitive optimal control solution because of the simplicity of the example and generally it's hard to find the* ***admissible*** *optimal control.*

Value Function, HJBE and the PDE Approach

Now the PDE approach to stochastic control focuses on applying the **dynamic programming principle**, setting up **value functions** and deriving **HJBE of the value functions** to solve the problem.

Let's assume that we are under **the Markovian setting** and the cost after time t sticking to control α with initial value condition $X_t = x$ is denoted as

$$J(t,x,\alpha) = \mathbb{E}\left[\int_t^T f(s,X_s,\alpha_s)\,ds + g(X_T)\Big| X_t = x\right] \tag{266}$$

(note that here X_s is the solution to the SDE for given control α) and let's denote $\mathscr{A}_t$ as the admissible set of controls α over time interval $[t,T]$ with measurability and integrability conditions

$$\mathbb{E}\int_t^T ||b(s,X_s,\alpha_s)||^2 + ||\sigma(s,X_s,\alpha_s)||^2\,ds < \infty \tag{267}$$

and the HJB value function is defined as

$$v(t,x) = \inf_{\alpha\in\mathscr{A}_t} J(t,x,\alpha) \tag{268}$$

the lowest possible cost with initial condition $X_t = x$ over all admissible controls. Since we are planning to find the HJBE that such value function is satisfying, it's natural to ask whether the value function is differentiable at all points and what conditions are needed such that a PDE for the value function can be constructed.

Remark. *We can also denote $J(t,x,\alpha) = \mathbb{E}\left[\int_t^T f(s,X_s^{t,x,\alpha},\alpha_s)\,ds + g(X_T^{t,x,\alpha})\right]$ and $X_s^{t,x,\alpha}$ denotes the value of the solution to the SDE at time s with fixed control α and initial value condition $X_t = x$. It's easy to see that those two definitions are equivalent.*

Example: Regularity Issues of Value Function

Let's look at an example where $d = 1$, i.e. X_t is 1-dimensional process with $A = [-1,1], \sigma = 0, f = 0, b(t,x,\alpha) = \alpha, g(x) = -x^2$. So now we know that

$$dX_t = \alpha_t\,dt \tag{269}$$

$$J(t,x,\alpha) = \mathbb{E}\left[-X_T^2\Big| X_t = x\right] \tag{270}$$

now conditioning on $X_t = x$, we know $X_s = x + \int_t^s \alpha_r \, dr$ so

$$J(t, x, \alpha) = -\left(x + \int_t^T \alpha_r \, dr\right)^2 \tag{271}$$

$$v(t, x) = \inf_{\alpha \in \mathscr{A}_t} \left\{ -\left(x + \int_t^T \alpha_r \, dr\right)^2 \right\} \tag{272}$$

it's clear that if $x \geq 0$ then since $\forall s \in [t, T], \alpha_s \in A = [-1, 1]$, the inf is attained when $\forall s \in [t, T], \alpha_s = 1$ and if $x \leq 0$ then the inf is attained when $\forall s \in [t, T], \alpha_s = -1$. So the value function is

$$v(t, x) = \begin{cases} -(x + T - t)^2 & x \geq 0 \\ -(x - T + t)^2 & x < 0 \end{cases} \tag{273}$$

which is continuous but not differentiable at 0 even under this extremely simple setting.

Example: Value Function as Convex Envelope

Consider another example where $d = k = 1, b = 0$ so X_t and the control α_t are still 1-dimensional and $\sigma(t, x, \alpha) = \alpha$ with $f = 0$, g continuous and bounded from above and $A = \mathbb{R}$. So now

$$dX_t = \alpha_t \, dB_t \tag{274}$$

$$J(t, x, \alpha) = \mathbb{E}\left[g(X_T) | X_t = x\right] \tag{275}$$

since we will be varying the time variable from t to $t + h$ a little bit for $h \to 0^+$ to set up a PDE for the value function (as shown later), it's natural to see that we would want to apply Ito formula for the value function v, so whether $v \in C^{1,2}$ is then a problem of our concern. However, in this example, we can show that **if $v \in C^{1,2}$, then v is independent of time t and is equal to the convex envelope g^{**} of g** (g^{**} is the double Frenchel conjugate of g, it can be proved that it's the convex envelope).

Now that $X_s|_{X_t = x} = x + \int_t^s \alpha_r \, dB_r$ and note that $\{\alpha_r, r \in [t, s]\} \in \mathscr{A}_t$ satisfies the integrability condition that $\mathbb{E} \int_t^T \alpha_s^2 \, ds < \infty$, it's obvious that $X_s|_{X_t = x}$ is a MG in s for any given control $\alpha \in \mathscr{A}_t$. Then we find

$$v(t, x) = \inf_{\alpha \in \mathscr{A}_t} \mathbb{E}\left[g(X_T) | X_t = x\right] \tag{276}$$

$$\geq \inf_{\alpha \in \mathscr{A}_t} \mathbb{E}\left[g^{**}(X_T) | X_t = x\right] \tag{277}$$

$$\geq \inf_{\alpha \in \mathscr{A}_t} g^{**}(\mathbb{E}(X_T | X_t = x)) \tag{278}$$

$$= \inf_{\alpha \in \mathscr{A}_t} g^{**}(x) \tag{279}$$

$$= g^{**}(x) \tag{280}$$

by applying Jensen's inequality, so v has to be larger than the convex envelope for $\forall x\in\mathbb{R}$.

For the other side, if $v\in C^{1,2}$, Ito formula holds and

$$v(t+h,X_{t+h}) = v(t,X_t) + \int_t^{t+h}\partial_t v(s,X_s)\,ds + \int_t^{t+h}\partial_x v(s,X_s)\,dX_s + \frac{1}{2}\int_t^{t+h}\partial_{xx}v(s,X_s)\,d\langle X,X\rangle_s \tag{281}$$

$$= v(t,X_t) + \int_t^{t+h}\left(\partial_t + \frac{\alpha_s^2}{2}\partial_{xx}\right)v(s,X_s)\,ds + \int_t^{t+h}\partial_x v(s,X_s)\cdot\alpha_s\,dB_s \tag{282}$$

assume that the last stochastic integral is a MG, we can find that

$$\mathbb{E}\left[v(t+h,X_{t+h})|X_t=x\right] = v(t,x) + \mathbb{E}\left[\int_t^{t+h}\left(\partial_t + \frac{\alpha_s^2}{2}\partial_{xx}\right)v(s,X_s)\,ds\Big|X_t=x\right] \tag{283}$$

by applying the property of value function that $\mathbb{E}\left[v(t+h,X_{t+h})|X_t=x\right]-v(t,x)\geq 0$ (which will be proved below), we get that

$$\forall h>0, \mathbb{E}\left[\int_t^{t+h}\left(\partial_t + \frac{\alpha_s^2}{2}\partial_{xx}\right)v(s,X_s)\,ds\Big|X_t=x\right]\geq 0 \tag{284}$$

dividing both sides by h and take $h\to 0^+$ to find

$$\forall (t,x,\alpha)\in[0,T]\times\mathbb{R}\times\mathbb{R}, \left(\partial_t + \frac{\alpha_t^2}{2}\partial_{xx}\right)v(t,x)\geq 0 \tag{285}$$

by taking $\alpha_t=0$ as the constant control, we find that $\partial_t v\geq 0$, so for each fixed $x\in\mathbb{R}$, v is always increasing w.r.t. time t. Also note that $\partial_{xx}v\geq 0$ must hold since otherwise we can always take α_t to be large enough such that the inequality above fails, so v has to be convex in x. By Fatou's lemma and continuity of g,

$$\forall 0\leq t<T, v(t,x)\leq \lim_{s\nearrow T} v(s,x) \tag{286}$$

$$= \lim_{s\nearrow T}\inf_{\alpha\in\mathscr{A}_s}\mathbb{E}\left[g(X_T)|X_s=x\right] \tag{287}$$

$$\leq \overline{\lim}_{s\nearrow T}\mathbb{E}\left[g(X_T)|X_s=x\right] \tag{288}$$

$$\leq \mathbb{E}\left[\overline{\lim}_{s\nearrow T}g(X_T^{s,x})\right] \tag{289}$$

$$= g(x) \tag{290}$$

where $X_T^{s,x}$ denotes the solution to the SDE with initial value condition $X_s=x$. So for any fixed time t, v is always a convex function dominated by g, $v(t,x)\leq g^{**}(x)$ by the maximality of convex envelope, and we conclude that

$$v(t,x) = g^{**}(x) \tag{291}$$

actually has nothing to do with t.

Remark. *The inequality $\mathbb{E}\left[v(t+h, X_{t+h})|X_t = x\right] - v(t,x) \geq 0$ we are using here is a natural property of the value function. The meaning is that since value function is already the optimal cost among all admissible controls based on the observation of the initial value condition $X_t = x$,* ***if one adopts the control that is the optimal control at time t in time interval $[t, t+h)$ but follows the optimal control at time $t+h$ in time interval $[t+h, T]$, then such strategy cannot be better than following the optimal control at time t in time interval $[t, T]$.*** *This would be explained in a later context.*

Remark. *This example shows that we can easily make up a "bad" value function. Let's pick g continuous and upper bounded on $\mathbb{R}$ with the convex envelope g^{**} being not C^2, then obviously $v(t,x)$ for this stochastic control problem would not be $C^{1,2}$. For example, consider $g(x) = -|x|$ then $g^{**} = g = v$. This is a type of control problems called* ***singular stochastic control problem****.*

However, as proved in the deterministic case for HJE, when $f = 0$ and g is Lipschitz, we can make sure that v is Lipschitz in x for fixed time $t \in [0,T]$ and when A is bounded one can also get some estimates on $|v(t,x) - v(\hat{t}, x)|$. Since **Lipschitz functions are almost everywhere differentiable**, this makes it possible for us to set up a PDE for the value function. We only prove the Lipschitz property here for simplicity.

Theorem 11. ***(Lipschitz Value Function in x)*** *When $f = 0$ and g is Lipschitz, the value function v is Lipschitz in x for fixed time $t \in [0,T]$.*

Proof. Fix time $t \in [0,T]$ and consider $\forall x, \hat{x} \in \mathbb{R}^d$, by the definition of value function, $\forall \varepsilon > 0, \exists \hat{\alpha}, v(t,\hat{x}) + \varepsilon \geq \mathbb{E}\left[g(X_T)|X_t = \hat{x}\right]$ with the X_t as the solution to the SDE with fixed control $\hat{\alpha}$

$$
\begin{aligned}
v(t,x) - v(t,\hat{x}) &\leq \inf_{\alpha \in \mathscr{A}_t} \mathbb{E}[g(X_T)|X_t = x] - \mathbb{E}\left[g(X_T)|X_t = \hat{x}\right] + \varepsilon && (292)\\
&\leq \mathbb{E}g(X_T^{t,x,\hat{\alpha}}) - \mathbb{E}g(X_T^{t,\hat{x},\hat{\alpha}}) + \varepsilon && (293)\\
&\leq Lips(g) \cdot \mathbb{E}|X_T^{t,x,\hat{\alpha}} - X_T^{t,\hat{x},\hat{\alpha}}| + \varepsilon && (294)
\end{aligned}
$$

let's consider $h(s) = \mathbb{E}|X_s^{t,x,\hat{\alpha}} - X_s^{t,\hat{x},\hat{\alpha}}|$, then

$$
\begin{aligned}
h(s) &= \mathbb{E}\left| x - \hat{x} + \int_t^s [b(r, X_r^{t,x,\hat{\alpha}}, \hat{\alpha}_r) - b(r, X_r^{t,\hat{x},\hat{\alpha}}, \hat{\alpha}_r)]\,dr + \int_t^s [\sigma(r, X_r^{t,x,\hat{\alpha}}, \hat{\alpha}_r) - \sigma(r, X_r^{t,\hat{x},\hat{\alpha}}, \hat{\alpha}_r)]\,dr \right| && (295)\\
&\leq |x-\hat{x}| + (Lips(b) + Lips(\sigma)) \cdot \mathbb{E}\int_t^s |X_r^{t,x,\hat{\alpha}} - X_r^{t,\hat{x},\hat{\alpha}}|\,dr && (296)\\
&= |x-\hat{x}| + (Lips(b) + Lips(\sigma)) \cdot \int_t^s h(r)\,dr && (297)
\end{aligned}
$$

since we have assumed that b, σ are both Lipschitz. By Grownwall's inequality,

$$
\forall s \in [t,T], h(s) \leq |x - \hat{x}| \cdot e^{C(s-t)} \tag{298}
$$

so we conclude that $v(t,x) - v(t,\hat{x}) \leq C \cdot h(T) + \varepsilon \leq C' \cdot |x - \hat{x}|$ which completes half of the proof. The other side can be done similarly. □

Dynamic Programming Principle (DPP)

Theorem 12. ***(Dynamic Programming Principle)*** *If the value function v is continuous, then for any initial value condition (t, x), and any stopping time τ that takes values in $[t, T]$,*

$$v(t,x) = \inf_{\alpha\in\mathscr{A}_t} \mathbb{E}\left[\int_t^\tau f(s, X_s^{t,x,\alpha}, \alpha_s)\,ds + v(\tau, X_\tau^{t,x,\alpha})\right] \tag{299}$$

Remark. *One may find that this theorem is an analogue of the optimality condition we have mentioned above for deterministic optimal control problem. The meaning of the theorem is that the best control α at time t can be found by minimizing the sum of two parts: (i): the contribution of cost in time interval $[t, \tau]$ sticking to control α (ii): the contribution of cost in time interval $[\tau, T]$ sticking to the optimal control at time τ. So this theorem is showing us* ***the time consistency condition for the value function****.*

Note that the difference of this property in deterministic control and stochastic control lies in the fact that: (i): we are taking inf of an expectation since the cost is actually random (ii): the deterministic perturbed time $t + h$ can be replaced by any stopping time that takes values in $[t, T]$, allowing us to have more freedom.

Proof. Notice that $\forall \alpha \in \mathscr{A}_t, v(t, x) \leq J(t, x, \alpha)$ so let us write $J(t, x, \alpha)$ in terms of the stopping time $\forall \tau \in \tau_{t,T}$ (where $\tau_{t,T}$ denotes the set of all stopping time that takes values in $[t, T]$)

$$\begin{aligned}
J(t,x,\alpha) &= \mathbb{E}\left[\int_t^T f(s, X_s, \alpha_s)\,ds + g(X_T)\Big| X_t = x\right] &(300)\\
&= \mathbb{E}\left[\mathbb{E}\left(\int_t^T f(s, X_s, \alpha_s)\,ds + g(X_T)\Big|\mathscr{F}_\tau\right)\Big| X_t = x\right] &(301)\\
&= \mathbb{E}\left[\int_t^\tau f(s, X_s, \alpha_s)\,ds + \mathbb{E}\left(\int_\tau^T f(s, X_s, \alpha_s)\,ds + g(X_T)\Big|\mathscr{F}_\tau\right)\Big| X_t = x\right] &(302)\\
&= \mathbb{E}\left[\int_t^\tau f(s, X_s, \alpha_s)\,ds + J(\tau, X_\tau, \alpha)\Big| X_t = x\right] &(303)
\end{aligned}$$

by tower property. Now replace the J inside the condition expectation with value function v to find

$$J(t,x,\alpha) \geq \mathbb{E}\left[\int_t^\tau f(s, X_s, \alpha_s)\,ds + v(\tau, X_\tau)\Big| X_t = x\right] \tag{304}$$

and take inf on both sides w.r.t. $\tau \in \tau_{t,T}$, take inf on both sides w.r.t. control $\alpha \in \mathscr{A}_t$ to find

$$v(t,x) \geq \inf_{\alpha\in\mathscr{A}_t} \inf_{\tau\in\tau_{t,T}} \mathbb{E}\left[\int_t^\tau f(s, X_s^{t,x,\alpha}, \alpha_s)\,ds + v(\tau, X_\tau^{t,x,\alpha})\right] \tag{305}$$

On the other hand, since value function is the inf of cost, $\forall \varepsilon > 0, \forall \tau \in \tau_{t,T}$, there exists **$\varepsilon$-optimal strategy** $\alpha^\varepsilon \in \mathscr{A}_\tau$ such that

$$v(\tau, X_\tau) + \varepsilon \geq J(\tau, X_\tau, \alpha^\varepsilon) \tag{306}$$

in order to argue that the value function has some upper bound, let's figure out the best control we are able to design so far. Now α^ε is nearly the best control at time τ, so we would expect to see that sticking to any current control $\alpha \in \mathscr{A}_t$ until stopping time τ and switch to the nearly best control α^ε after time τ would be a good strategy (Note that here we are **switching to the best strategy at time** τ since the admissible control $\alpha^\varepsilon \in \mathscr{A}_\tau$ should satisfy the measurability condition under the Markov setting that $\alpha^\varepsilon = \alpha^\varepsilon(\tau, X_\tau) \in \mathscr{F}_\tau$ so there's no way to know α^ε before time τ). So we construct

$$\hat{\alpha}_s = \begin{cases} \alpha_s & s \in [t, \tau] \\ \alpha_s^\varepsilon & s \in [\tau, T] \end{cases} \in \mathscr{A}_t \tag{307}$$

and

$$\forall \alpha \in \mathscr{A}_t, v(t,x) \le J(t,x,\hat{\alpha}) \tag{308}$$

$$= \mathbb{E}\left[\int_t^T f(s, X_s^{\hat{\alpha}}, \hat{\alpha}_s)\, ds + g(X_T^{\hat{\alpha}}) \Big| X_t = x\right] \tag{309}$$

$$= \mathbb{E}\left[\int_t^\tau f(s, X_s^{\alpha}, \alpha_s)\, ds + \int_\tau^T f(s, X_s^{\hat{\alpha}}, \hat{\alpha}_s)\, ds + g(X_T^{\hat{\alpha}}) \Big| X_t = x\right] \tag{310}$$

$$= \mathbb{E}\left[\int_t^\tau f(s, X_s^{\alpha}, \alpha_s)\, ds + J(\tau, X_\tau, \hat{\alpha}) \Big| X_t = x\right] \tag{311}$$

$$\le \mathbb{E}\left[\int_t^\tau f(s, X_s^{\alpha}, \alpha_s)\, ds + v(\tau, X_\tau) \Big| X_t = x\right] + \varepsilon \tag{312}$$

by first taking the sup w.r.t. $\tau \in \tau_{t,T}$ on both sides and then the inf on both sides w.r.t. control $\alpha \in \mathscr{A}_t$ to find

$$v(t,x) \le \inf_{\alpha \in \mathscr{A}_t} \sup_{\tau \in \tau_{t,T}} \mathbb{E}\left[\int_t^\tau f(s, X_s^{t,x,\alpha}, \alpha_s)\, ds + v(\tau, X_\tau^{t,x,\alpha})\right] \tag{313}$$

Combining two inequalities to see the DPP

$$v(t,x) = \inf_{\alpha \in \mathscr{A}_t} \sup_{\tau \in \tau_{t,T}} \mathbb{E}\left[\int_t^\tau f(s, X_s^{t,x,\alpha}, \alpha_s)\, ds + v(\tau, X_\tau^{t,x,\alpha})\right] = \inf_{\alpha \in \mathscr{A}_t} \inf_{\tau \in \tau_{t,T}} \mathbb{E}\left[\int_t^\tau f(s, X_s^{t,x,\alpha}, \alpha_s)\, ds + v(\tau, X_\tau^{t,x,\alpha})\right] \tag{314}$$

□

Remark. *We have actually shown that* ***the selection of stopping time has no impact on the value function****, so any stopping time $\tau \in \tau_{t,T}$ works. This is because we are first fixing the stopping time and then select a good enough control,* ***the inf w.r.t. control α has already taken the stopping time into consideration!***

HJBE of Stochastic Control Problem

Taking the stopping time in DPP as the trivial one $\tau = t+h$ such that $h>0, t+h\leq T$, one would get the following HJBE for the stochastic control problem. Let's denote L^α as the **infinitesimal generator** of the diffusion process X_t for fixed control α. Then from stochastic calculus, we know that

$$L^\alpha f(x) = b(t,x,\alpha)\cdot\nabla_x f(x) + \frac{1}{2}Tr\left(\sigma(t,x,\alpha)\cdot\sigma^T(t,x,\alpha)\cdot\nabla_x^2 f(x)\right) \tag{315}$$

where $b\in\mathbb{R}^d, \sigma\in\mathbb{R}^{d\times m}$ are drift and diffusion coefficients of the dynamics, $\nabla_x f$ is the gradient of f w.r.t. variable x and $\nabla_x^2 f$ is the Hessian of f w.r.t. variable x.

Theorem 13. ***(HJBE of Stochastic Control Problem)*** *Assume that $v\in C^{1,2}([0,T]\times\mathbb{R}^d)$ and $f\in C([0,T]\times\mathbb{R}^d\times A)$ for each fixed control $\alpha\in\mathscr{A}$ and assume the existence of the optimal control $\alpha^*\in\mathscr{A}$. Then*

$$\forall (t,x)\in[0,T]\times\mathbb{R}^d, \partial_t v(t,x) + \inf_{\alpha\in\mathscr{A}}\{L^\alpha v(t,x) + f(t,x,\alpha)\} = 0 \tag{316}$$

Proof. By taking the stopping time in DPP as the trivial one $\tau = t+h$ such that $h>0, t+h\leq T$, we find that

$$v(t,x) = \inf_{\alpha\in\mathscr{A}_t}\mathbb{E}\left[\int_t^{t+h} f(s,X_s^{t,x,\alpha},\alpha_s)\,ds + v(t+h, X_{t+h}^{t,x,\alpha})\right] \tag{317}$$

apply Ito formula to see

$$v(t+h, X_{t+h}^{t,x,\alpha}) = v(t, X_t^{t,x,\alpha}) + \int_t^{t+h}\partial_t v(s, X_s^{t,x,\alpha})\,ds + \int_t^{t+h}\partial_x v(s,X_s^{t,x,\alpha})\,dX_s^{t,x,\alpha} \tag{318}$$

$$+\frac{1}{2}\int_t^{t+h}\partial_{xx}v(s,X_s^{t,x,\alpha})\,d\langle X^{t,x,\alpha}, X^{t,x,\alpha}\rangle_s \tag{319}$$

$$= v(t,x) + \int_t^{t+h}\left(\partial_t + L^\alpha\right)v(s,X_s^{t,x,\alpha})\,ds + \int_t^{t+h}\partial_x v(s,X_s^{t,x,\alpha})\cdot\sigma(s,X_s^{t,x,\alpha},\alpha_s)\,dB_s \tag{320}$$

assume that the last stochastic integral is a MG, we can find that

$$\mathbb{E}\left[v(t+h,X_{t+h})|X_t = x\right] = v(t,x) + \mathbb{E}\left[\int_t^{t+h}\left(\partial_t + L^\alpha\right)v(s,X_s^\alpha)\,ds\Big|X_t = x\right] \tag{321}$$

by noticing that from DPP we have

$$\forall\alpha\in\mathscr{A}_t, v(t,x)\leq\mathbb{E}\left[\int_t^{t+h} f(s,X_s^{t,x,\alpha},\alpha_s)\,ds + v(t+h,X_{t+h}^{t,x,\alpha})\right] \tag{322}$$

$$= v(t,x) + \mathbb{E}\left[\int_t^{t+h}\left(\partial_t + L^\alpha\right)v(s,X_s^\alpha) + f(s,X_s^\alpha,\alpha_s)\,ds\Big|X_t = x\right] \tag{323}$$

dividing both sides by h and apply the intermediate value theorem for integral, we get (note that the diffusion process

X_t is chosen as the version with continuous sample path and f is continuous)

$$\forall \alpha \in \mathscr{A}_t, 0 \leq (\partial_t + L^\alpha)v(t,x) + f(t,x,\alpha) \tag{324}$$

taking inf on both sides to see

$$\inf_{\alpha \in \mathscr{A}_t} \{(\partial_t + L^\alpha)v(t,x) + f(t,x,\alpha)\} \geq 0 \tag{325}$$

The equality directly comes from the assumption that the optimal control $\alpha^* \in \mathscr{A}_t$ exists and can attain the inf in the value function. As a result, the inequality in DPP becomes equality and

$$(\partial_t + L^{\alpha^*})v(t,x) + f(t,x,\alpha^*) = 0 \tag{326}$$

that's why we get the HJBE

$$\inf_{\alpha \in \mathscr{A}_t} \{(\partial_t + L^\alpha)v(t,x) + f(t,x,\alpha)\} = 0 \tag{327}$$

□

Remark. *Now one can see why the argument we have made in the example above that* $\mathbb{E}\left[v(t+h, X_{t+h})|X_t = x\right] \geq v(t,x)$ *is true. This is just a simple corollary of the DPP.*

Remark. *For maximization problem, just change the inf w.r.t. control in DPP into sup and DPP still holds, so the HJBE now becomes*

$$\forall (t,x) \in [0,T] \times \mathbb{R}^d, \partial_t v(t,x) + \sup_{\alpha \in \mathscr{A}} \{L^\alpha v(t,x) + f(t,x,\alpha)\} = 0 \tag{328}$$

Verification Theorem

Now that we know the value function satisfies the HJBE, and we wonder whether solving the HJBE necessarily gives the value function. In other words, we want to see that HJBE actually **characterizes** the value function. In such case, by solving the HJBE we would know the value function immediately, which is the key to the PDE approach of solving stochastic control problems. The condition is given by the following verification theorem in the Markovian case.

Theorem 14. ***(Verification Theorem for Finite Horizon Case)*** *Let $w \in C^{1,2}([0,T]\times\mathbb{R}^d)\cap C([0,T]\cap\mathbb{R}^d)$ with growth condition*

$$\exists C>0, \forall (t,x)\in[0,T]\times\mathbb{R}^d, |w(t,x)|\le C(1+||x||^2) \tag{329}$$

now if

$$\begin{cases}\forall (t,x)\in[0,T]\times\mathbb{R}^d, -\partial_t w(t,x)-\inf_{\alpha\in\mathscr{A}}\{L^\alpha w(t,x)+f(t,x,\alpha)\}\le 0\\ \forall x\in\mathbb{R}^d, w(T,x)\le g(x)\end{cases} \tag{330}$$

then $w\le v$ for value function v.

Moreover, if $\forall x\in\mathbb{R}^d, w(T,x)=g(x)$ and exists measurable $\hat{\alpha}(t,x):[0,T]\times\mathbb{R}^d\to A$ such that

$$\forall (t,x)\in[0,T]\times\mathbb{R}^d, -\partial_t w(t,x)-\inf_{\alpha\in\mathscr{A}}\{L^\alpha w(t,x)+f(t,x,\alpha)\} = -\partial_t w(t,x)-[L^{\hat{\alpha}(t,x)}w(t,x)+f(t,x,\hat{\alpha}(t,x))]=0 \tag{331}$$

and $\hat{\alpha}(s,X_s^{t,x,\hat{\alpha}(t,x)})\in\mathscr{A}_t$, then $w=v$ for value function v and $\hat{\alpha}$ is the optimal Markovian control.

Proof. Consider applying Ito formula for $w(s\wedge\tau, X^{t,x,\alpha}_{s\wedge\tau})$, $\tau\in\tau_{t,\infty}$ is any stopping time that takes value in $[t,\infty)$

$$w(s\wedge\tau, X^{t,x,\alpha}_{s\wedge\tau}) = w(t,x)+\int_t^{s\wedge\tau}\partial_t w(u,X_u^{t,x,\alpha})\,du+\int_t^{s\wedge\tau}\partial_x w(u,X_u^{t,x,\alpha})\,dX_u^{t,x,\alpha} \tag{332}$$

$$+\frac{1}{2}\int_t^{s\wedge\tau}\partial_{xx}w(u,X_u^{t,x,\alpha})\,d\langle X^{t,x,\alpha},X^{t,x,\alpha}\rangle_u \tag{333}$$

$$=w(t,x)+\int_t^{s\wedge\tau}(\partial_t+L^\alpha)w(u,X_u^{t,x,\alpha})\,du+\int_t^{s\wedge\tau}\sigma(u,X_u^{t,x,\alpha},\alpha_u)\cdot\partial_x w(u,X_u^{t,x,\alpha})\,dB_u \tag{334}$$

now in order to bound the last stochastic integral term, choose the stopping time τ as

$$\tau_n=\inf\left\{s\ge t:\int_t^s[\sigma(u,X_u^{t,x,\alpha},\alpha_u)\cdot\partial_x w(u,X_u^{t,x,\alpha})]^2\,du\ge n\right\} \tag{335}$$

when the quadratic variation of the stochastic integral exceeds n we stop immediately. Such $\tau_n\nearrow\infty$ $(n\to\infty)$ and it's obvious that this series of stopping time reduces the local MG, so we conclude that now $\int_t^{s\wedge\tau_n}\sigma(u,X_u^{t,x,\alpha},\alpha_u)\cdot\partial_x w(u,X_u^{t,x,\alpha})\,dB_u$ is a U.I. MG in s on time interval $[t,T]$.

Take expectation on both sides to get rid of the stochastic integral term

$$\mathbb{E}w(s \wedge \tau_n, X^{t,x,\alpha}_{s\wedge\tau_n}) = w(t,x) + \mathbb{E}\int_t^{s\wedge\tau_n} (\partial_t + L^\alpha) w(u, X^{t,x,\alpha}_u)\, du \tag{336}$$

$$\geq w(t,x) - \mathbb{E}\int_t^{s\wedge\tau_n} f(u, X^{t,x,\alpha}_u, \alpha_u)\, du \tag{337}$$

because of the inequality condition and the result we have derived above shall hold for $\forall \alpha \in \mathscr{A}_t$. Notice that

$$\left|\int_t^{s\wedge\tau_n} f(u, X^{t,x,\alpha}_u, \alpha_u)\, du\right| \leq \int_t^{s\wedge\tau_n} \left|f(u, X^{t,x,\alpha}_u, \alpha_u)\right|\, du \tag{338}$$

$$\leq \int_t^T \left|f(u, X^{t,x,\alpha}_u, \alpha_u)\right|\, du \tag{339}$$

which is integrable (a natural assumption) and under the integrability condition of admissible control that $\mathbb{E}\int_0^T ||b(t,0,\alpha_t)||^2 + ||\sigma(t,0,\alpha_t)||^2\, dt < \infty$, we have $\mathbb{E}\sup_{s\in[t,T]} ||X^{t,x}_s||^2 < \infty$, so

$$\mathbb{E}w(s\wedge\tau_n, X^{t,x,\alpha}_{s\wedge\tau_n}) \leq C[1 + \mathbb{E}(X^{t,x,\alpha}_{s\wedge\tau_n})^2] \leq C\left[1 + \mathbb{E}\left(\sup_{s\in[t,T]} X^{t,x,\alpha}_s\right)^2\right] < \infty \tag{340}$$

by applying the growth condition on w. So now set $n \to \infty$ to find that

$$\mathbb{E}w(s, X^{t,x,\alpha}_s) \geq w(t,x) - \mathbb{E}\int_t^s f(u, X^{t,x,\alpha}_u, \alpha_u)\, du \tag{341}$$

by dominated convergence theorem. Set $s \to T^-$ to find that

$$\mathbb{E}g(X^{t,x,\alpha}_T) \geq \mathbb{E}w(T, X^{t,x,\alpha}_T) \geq w(t,x) - \mathbb{E}\int_t^T f(u, X^{t,x,\alpha}_u, \alpha_u)\, du \tag{342}$$

by dominated convergence theorem again. Now take inf w.r.t. the control $\alpha \in \mathscr{A}_t$ to get

$$v(t,x) \geq w(t,x) \tag{343}$$

When $w(T,x) = g(x)$ and exists $\hat{\alpha}(t,x)$, after removing the stochastic integral term (by similar stopping time argument as above), we get

$$\mathbb{E}w(s, X^{t,x,\hat{\alpha}}_s) = w(t,x) + \mathbb{E}\int_t^s (\partial_t + L^{\hat{\alpha}(t,x)}) w(u, X^{t,x,\hat{\alpha}}_u)\, du \tag{344}$$

$$= w(t,x) + \mathbb{E}\int_t^s f(u, X^{t,x,\hat{\alpha}}_u, \hat{\alpha}_u)\, du \tag{345}$$

by the definition of $\hat{\alpha}(t,x)$ as the control that achieves the inf in HJBE for all pair (t,x). By setting $s\to T^-$ again,

$$\mathbb{E}g(X_T^{t,x,\hat{\alpha}}) = \mathbb{E}w(T,X_T^{t,x,\hat{\alpha}}) = w(t,x) - \mathbb{E}\int_t^T f(u,X_u^{t,x,\hat{\alpha}},\hat{\alpha}_u)\,du \tag{346}$$

by dominated convergence theorem and one might find that the terms relevant with g,f give the definition of $J(t,x,\hat{\alpha}(t,x))$, so we see that

$$J(t,x,\hat{\alpha}(t,x)) = w(t,x) \tag{347}$$

naturally $w(t,x)\geq v(t,x)$ and the theorem is proved. □

Remark. *In particular, when the control space $\mathscr{A}=\{a_0\}$ only contains one single admissible control, HJBE turns into*

$$\begin{cases}\partial_t w(t,x) + L^\alpha w(t,x) + f(t,x,a_0) = 0\\ w(T,x) = g(x)\end{cases} \tag{348}$$

and the value function is

$$v(t,x) = J(t,x,a_0) = \mathbb{E}\left[\int_t^T f(s,X_s^{t,x,a_0},a_0)\,ds + g(X_T^{t,x,a_0})\right] \tag{349}$$

and the verification theorem is actually just the ***Feynman-Kac formula*** *since it tells us that the $C^{1,2}$ solution to such PDE is characterized by*

$$w(t,x) = \mathbb{E}\left[\int_t^T f(s,X_s^{a_0},a_0)\,ds + g(X_T^{a_0})\middle| X_t^{a_0}=x\right] = v(t,x) \tag{350}$$

Remark. *The time we use verification theorem is typically after solving the HJBE. We* ***check that the solution is $C^{1,2}$ has growth condition and check by implicit function theorem that the optimal control $\hat{\alpha}$ that minimizes the inf in HJBE can be written as a function of*** *(t,x). Then verification theorem tells us that such solution must be the value function.*

Stochastic Control Problem: PDE Approach

Now that we have described the main tools for the PDE approach to solve stochastic control problems. However, our description is not very organized since we have made a lot of different assumptions within the discussions. Let's collect all non-trivial assumptions we have made so far and present a systematic way for the PDE approach.

- Assume the value function $v \in C^{1,2}$ (in order to apply Ito's formula)
- Assume that the stochastic integral in the calculation is a true MG (in order to ignore it after taking expectation)
- Assume that the optimal control $\alpha^* \in \mathscr{A}$ always exists (in order to turn the inequality into equality in HJBE)
- Assume that for each initial value pair (t, x), there always exists $\hat{\alpha} = \hat{\alpha}(t, x)$ that minimizes $L^\alpha v(t, x) + f(t, x, \alpha)$ (condition of verification theorem)
- Solve the HJBE to get the value function, check the conditions of the **verification theorem** ($C^{1,2}$ solution, growth condition, existence of $\hat{\alpha}(t, x)$)
- After solving out the optimal control, check that it's admissible, the value function is $C^{1,2}$, and the local MG is actually a MG
- Note that $v \in C^{1,2}$ can also be checked by applying **uniform ellipticity** (lowest eigenvalue of $\sigma\sigma^T$ bounded away from 0, $\exists C > 0, \forall x, y \in \mathbb{R}^d, \alpha \in \mathscr{A}, y^T \sigma(x, \alpha)\sigma^T(x, \alpha)y \geq C||y||^2$)
- The existence of the minimizer $\hat{\alpha} = \hat{\alpha}(t, x)$ can be checked through convex analysis and **implicit function theorem**

Infinite Horizon Case

The infinite horizon case is actually the same as what we have done for HJE, the deterministic optimal control problem. For the purpose of completeness, we state it again here. The difference in formulation is the introduction of the **discount factor** $\beta > 0$ and the fact that no terminal reward exists. The expected cost after time t following control α is

$$J(t,x,\alpha) = \mathbb{E}\left[\int_t^\infty e^{-\beta s} f(s, X_s^{t,x,\alpha}, \alpha_s)\, ds\right] \tag{351}$$

the introduction of discount factor is to ensure that the integral will be finite for a general class of f. However, by changing variables $u = s - t$ and **assuming that f is time-homogeneous** ($f(s, X_s^{t,x,\alpha}, \alpha_s) = f(X_s^{t,x,\alpha}, \alpha_s)$) one might find that

$$\begin{aligned} J(t,x,\alpha) &= e^{-\beta t} \cdot \mathbb{E}\left[\int_0^\infty e^{-\beta s} f(X_{s+t}^{t,x,\alpha}, \alpha_{s+t})\, ds\right] && (352)\\ &= e^{-\beta t} \cdot \mathbb{E}\left[\int_0^\infty e^{-\beta s} f(X_s^{0,x,\alpha}, \alpha_s)\, ds\right] && (353)\\ &= e^{-\beta t} \cdot J(0,x,\alpha) && (354) \end{aligned}$$

under the Markovian setting. As a result, we remove the time variable for simplicity and consider a slightly different definition that

$$J(x,\alpha) \overset{def}{=} J(0,x,\alpha) = \mathbb{E}\left[\int_0^\infty e^{-\beta s} f(X_s^{0,x,\alpha}, \alpha_s)\, ds\right] \tag{355}$$

as a result, the value function is defined as

$$v(x) \overset{def}{=} \inf_{\alpha \in \mathscr{A}} J(x,\alpha) \tag{356}$$

independent of time and the connection of this value function with the previously defined value function is that

$$v(x) = v(0,x) = e^{\beta t} \cdot v(t,x) \tag{357}$$

Remark. *The β can be understood as the opposite of the continuous-time interest rate. Under the condition that the discount factor has the formulation as $e^{-\beta t}$ and the fact that f is time-homogeneous, we can eliminate the time variable in t (failure in either assumption would cause mistake in doing so).*

The reason we mention the connection between the value function that only works for the special infinite horizon case and that for the general case is that we would still be able to apply the general results to get the HJBE.

Theorem 15. ***(HJBE of Stochastic Control Problem in Infinite Horizon Case)*** *Assume that $v \in C^2(\mathbb{R}^d)$*

and $f \in C(\mathbb{R}^d \times A)$ for each fixed control $\alpha \in \mathscr{A}$ and assume the existence of the optimal control $\alpha^* \in \mathscr{A}$. Then

$$\forall x \in \mathbb{R}^d, -\beta v(x) + \inf_{\alpha \in \mathscr{A}} \{L^\alpha v(x) + f(x,\alpha)\} = 0 \tag{358}$$

Proof. Denote the general value function as $u(t,x)$ so now $v(x) = u(0,x) = e^{\beta t} \cdot u(t,x)$. Since $\partial_t u(t,x) = -\beta e^{-\beta t} \cdot v(x), L^\alpha u(t,x) = e^{-\beta t} \cdot L^\alpha v(x)$, apply the HJBE for $u(t,x)$ to know that

$$\forall t > 0, x \in \mathbb{R}^d, , -\beta e^{-\beta t} \cdot v(x) + \inf_{\alpha \in \mathscr{A}} \left\{ e^{-\beta t} \cdot L^\alpha v(x) + f(x,\alpha) \right\} = 0 \tag{359}$$

set $t = 0$ to get

$$\forall x \in \mathbb{R}^d, -\beta v(x) + \inf_{\alpha \in \mathscr{A}} \{L^\alpha v(x) + f(x,\alpha)\} = 0 \tag{360}$$

$\square$

For the sake of completeness, let's refer to the verification theorem in the infinite horizon case mentioned above. The statements are a little bit different but the thought of the proof is the same.

Theorem 16. ***(Verification Theorem for Infinite Horizon Case)*** *Let $w \in C^2(\mathbb{R}^d)$ with growth condition*

$$\exists C > 0, \forall x \in \mathbb{R}^d, |w(x)| \le C(1 + ||x||^2) \tag{361}$$

now if

$$\begin{cases} \forall x \in \mathbb{R}^d, \beta w(x) - \inf_{\alpha \in \mathscr{A}} \{L^\alpha w(x) + f(x,\alpha)\} \le 0 \\ \forall x \in \mathbb{R}^d, \forall \alpha \in \mathscr{A}, \underline{\lim}_{T \to \infty} e^{-\beta T} \cdot \mathbb{E} w(X_T^{0,x,\alpha}) \le 0 \end{cases} \tag{362}$$

then $w \le v$ for value function v.

Moreover, if there exists measurable $\hat{\alpha}(x) : \mathbb{R}^d \to A$ such that

$$\forall x \in \mathbb{R}^d, \beta w(x) - \inf_{\alpha \in \mathscr{A}} \{L^\alpha w(x) + f(x,\alpha)\} = \beta w(x) - [L^{\hat{\alpha}(x)} w(x) + f(x, \hat{\alpha}(x))] = 0 \tag{363}$$

and $\hat{\alpha}(X_s^{0,x,\hat{\alpha}(x)}) \in \mathscr{A}, \overline{\lim}_{T \to \infty} e^{-\beta T} \cdot \mathbb{E} w(X_T^{0,x,\hat{\alpha}(x)}) \ge 0$, then $w = v$ for value function v and $\hat{\alpha}$ is the optimal Markovian control.

Proof. We just list the sketch of the proof here. The proof is almost the same as that for finite horizon case. Just consider applying Ito formula for $e^{-\beta(T \wedge \tau_n)} w(X_{T \wedge \tau_n}^{0,x})$ with discount factor where τ_n is still the stopping time that reduces the local MG and T is any positive number, take expectation on both sides and use dominated convergence theorem (set $T \to \infty$) to conclude. $\square$

Remark. *The reason that two extra conditions $\underline{\lim}_{T \to \infty} e^{-\beta T} \cdot \mathbb{E} w(X_T^{0,x,\alpha}) \le 0$ and $\overline{\lim}_{T \to \infty} e^{-\beta T} \cdot \mathbb{E} w(X_T^{0,x,\hat{\alpha}(x)}) \ge 0$ appear is due to the difference that in infinite horizon case we don't have the trivial bound that $s \wedge \tau_n \le T$ for a fixed* *time limit T and we don't have terminal costs.*

In brief, besides ***checking that the solution to HJBE is C^2, has the growth condition and the existence of $\hat{\alpha}(x)$, we also have to check other two mild conditions on the tail growth rate of the expectation of the solution compared to the discount factor****. When the admissible control space $\mathscr{A} = \{a_0\}$, note again that verification theorem is just the* ***Feynman-Kac formula****.*

Remark. *For maximization problems, in the finite horizon case still check the same conditions while in the infinite horizon case, the two extra conditions become*

$$\overline{\lim}_{T\to\infty} e^{-\beta T} \cdot \mathbb{E}w(X_T^{0,x,\alpha}) \geq 0, \underline{\lim}_{T\to\infty} e^{-\beta T} \cdot \mathbb{E}w(X_T^{0,x,\hat{\alpha}(x)}) \leq 0 \tag{364}$$

Example: Merton Problem with Consumption

Let's consider the Merton problem in infinite time horizon $t \in [0, \infty)$. Now we have one riskless asset (say, bond) with price P_t^0 and one risky asset (say, stock) with price P_t^1 at time t following the dynamics

$$\begin{cases} dP_t^0 = rP_t^0 \, dt \\ P_0^0 = 1 \\ dP_t^1 = P_t^1 \cdot (\mu \, dt + \sigma \, dB_t) \\ P_0^1 = p \end{cases} \tag{365}$$

with r, μ, σ as constants and $r < \mu$ (nontrivial case). Now one would always invest α_t of his total wealth into the stock at time t and c_t as consumption rate (integrates to the amount of consumption), investing all remaining wealth into the bond. Denote X_t as his total wealth at time t, one would find that if one sticks to (α_t, c_t) from time t to time $t + h$ for $h \to 0^+$, an infinitesimal time increment, one would get

$$X_{t+h} = \alpha_t X_t \cdot \frac{P_{t+h}^1}{P_t^1} + (X_t - c_t h - \alpha_t X_t) \cdot e^{rh} \tag{366}$$

wealth at time $t + h$ (the consumption amount is $c_t h$ since the length of the time interval is h). Now write it as a SDE for X_t to get

$$dX_t = \alpha_t X_t \frac{dP_t^1}{P_t^1} + (1 - \alpha_t) r X_t \, dt - c_t \, dt \tag{367}$$

and plug in the SDE for stock price to get

$$dX_t = \alpha_t X_t(\mu \, dt + \sigma \, dB_t) + (1 - \alpha_t) r X_t \, dt - c_t \, dt \tag{368}$$

$$dX_t = (X_t[\alpha_t \mu + r(1 - \alpha_t)] - c_t) \, dt + \alpha_t \sigma X_t \, dB_t \tag{369}$$

As a result, the control is a pair $(\alpha_t, c_t) \in \mathscr{A}$ for $A = \mathbb{R} \times [0, \infty)$ (allow the shorting of stocks) and the admissible set is $\mathscr{A} = \left\{(\alpha_t, c_t) : (0, \infty) \to A : \int_0^\infty |\alpha_t|^2 + c_t \, dt < \infty \ a.s., (\alpha_t, c_t) = (\alpha_t(t, X_t), c_t(t, X_t))\right\}$ under the Markovian setting. The integrability condition ensures that the SDE for X_t has unique strong solution. Now our goal is to maximize the expected utility that depends on the consumption rate with given discount factor $\beta > 0$

$$\mathbb{E}\left[\int_0^\infty e^{-\beta t} U(c_t) \, dt\right] \tag{370}$$

for the CRRA (constant relative risk aversion) utility function $U(x) = \frac{x^{1-\gamma}}{1-\gamma}$ ($\gamma \in (0, 1)$), so the value function is formed as

$$v(x) = \sup_{(\alpha, c) \in \mathscr{A}} \mathbb{E}\left[\int_0^\infty e^{-\beta t} U(c_t) \, dt \middle| X_0 = x\right] \tag{371}$$

since we are in the infinite horizon case and the utility function is time-homogeneous.

By previous proved theorem, the HJBE is now

$$-\beta v(x) + \sup_{(\alpha,c)\in\mathscr{A}} \{L^{\alpha,c}v(x) + f(x,\alpha,c)\} = 0 \tag{372}$$

plug in the expressions to find

$$\begin{cases} L^{\alpha,c}v(x) = (x[\alpha\mu + r(1-\alpha)] - c)\cdot v'(x) + \frac{\alpha^2\sigma^2x^2}{2}\cdot v''(x) \\ f(x,\alpha,c) = U(c) \end{cases} \tag{373}$$

so the **HJBE** is given by

$$-\beta v + \sup_{(\alpha,c)\in\mathscr{A}} \left\{ \frac{\alpha^2\sigma^2x^2}{2}\cdot v'' + [x(\alpha\mu + r - \alpha r) - c]\cdot v' + U(c) \right\} = 0 \tag{374}$$

To solve this HJBE, of course let's first get rid of the sup by solving an optimization problem w.r.t. (α, c) that

$$\max_{\alpha,c}\ Q(x,\alpha,c) = \frac{\alpha^2\sigma^2x^2}{2}\cdot v'' + [x(\alpha\mu + r - \alpha r) - c]\cdot v' + U(c) \tag{375}$$

and compute the partials

$$\begin{cases} \frac{\partial Q}{\partial \alpha} = x\mu v' - xrv' + \sigma^2x^2\alpha v'' \\ \frac{\partial Q}{\partial c} = U'(c) - v' \end{cases} \tag{376}$$

to solve out the optimal α^*, c^* as the function of $v'v''$

$$\begin{cases} \alpha^* = \frac{(r-\mu)v'}{\sigma^2xv''} \\ c^* = (U')^{-1}(v') = (v')^{-\frac{1}{\gamma}} \end{cases} \tag{377}$$

and plug once again back into the HJBE to get the ODE

$$rxv' - \frac{(\mu-r)^2(v')^2}{2\sigma^2v''} - \beta v + \frac{\gamma}{1-\gamma}(v')^{\frac{\gamma-1}{\gamma}} = 0 \tag{378}$$

We try the **ansatz** that $v(x) = \kappa U(x) = \kappa\frac{x^{1-\gamma}}{1-\gamma}$ to get

$$\kappa = \left(\frac{\beta - r(1-\gamma)}{\gamma} - \frac{(\mu-r)^2(1-\gamma)}{2\sigma^2\gamma^2}\right)^{-\gamma} \tag{379}$$

so now the optimal control is given by

$$\begin{cases} \alpha^* = \frac{(r-\mu)v'}{\sigma^2 x v''} = \frac{\mu - r}{\gamma\sigma^2} \\ c^* = (U')^{-1}(v') = (v')^{-\frac{1}{\gamma}} = \left(\frac{\beta - r(1-\gamma)}{\gamma} - \frac{(\mu-r)^2(1-\gamma)}{2\sigma^2\gamma^2}\right) \cdot x \end{cases} \tag{380}$$

the optimal choice is to always consume and invest in risky asset a fixed proportion of total wealth.

Verification of Merton Problem with Consumption

Now the calculation part comes to an end but there's still some verification to do. As stated above, the first thing to do is to verify the condition of the verification theorem. Now the solution we have found is

$$v(x) = \kappa \frac{x^{1-\gamma}}{1-\gamma} \tag{381}$$

so it's in $C^2(\mathbb{R})$, satisfies the growth condition, $\alpha^* = \alpha^*(x), c^* = c^*(x)$ exist and are well-defined, the condition $\overline{\lim}_{T\to\infty} e^{-\beta T} \cdot \mathbb{E}v(X_T^{0,x,(\alpha,c)}) \geq 0$ is satisfied due to non-negativity. Consider the SDE

$$\begin{cases} dX_t = (X_t[\alpha_t^*\mu + r(1-\alpha_t^*)] - c_t^*)\, dt + \alpha_t^*\sigma X_t\, dB_t \\ X_0 = x \end{cases} \tag{382}$$

then it's easy to see that this is a geometric BM (by denoting $c^* = \nu x$, so $c_t^* = \nu X_t$), so the solution exists and is unique

$$X_T^{0,x,(\alpha^*,c^*)} = x \cdot e^{\left[\alpha^*\mu + r(1-\alpha^*) - \nu - \frac{\sigma^2(\alpha^*)^2}{2}\right]T + \alpha^*\sigma B_T} \tag{383}$$

some calculations show

$$\mathbb{E}v\left(X_T^{0,x,(\alpha^*,c^*)}\right) = \frac{\kappa}{1-\gamma} \cdot \mathbb{E}\left(X_T^{0,x,(\alpha^*,c^*)}\right)^{1-\gamma} \tag{384}$$

$$= \kappa \frac{x^{1-\gamma}}{1-\gamma} \cdot e^{\left[\alpha^*\mu + r(1-\alpha^*) - \nu - \frac{\sigma^2(\alpha^*)^2}{2}\right](1-\gamma)T + \frac{(\alpha^*)^2\sigma^2(1-\gamma)^2}{2}T} \tag{385}$$

$$= \kappa \frac{x^{1-\gamma}}{1-\gamma} \cdot e^{(1-\gamma)\left[\frac{(\mu-r)^2}{2\gamma^2\sigma^2} + \frac{r-\beta}{\gamma}\right]T} \tag{386}$$

$$\tag{387}$$

so

$$\underline{\lim}_{T\to\infty} e^{-\beta T} \cdot \mathbb{E}v\left(X_T^{0,x,(\alpha^*,c^*)}\right) = \kappa \frac{x^{1-\gamma}}{1-\gamma} \cdot \underline{\lim}_{T\to\infty} e^{\left((1-\gamma)\left[\frac{(\mu-r)^2}{2\gamma^2\sigma^2} + \frac{r-\beta}{\gamma}\right] - \beta\right)T} \tag{388}$$

in order to make sure that this expression is non-positive in order to make the verification theorem true, we have to **assume that** $\beta > (1-\gamma)\left[\frac{(\mu-r)^2}{2\gamma^2\sigma^2} + \frac{r-\beta}{\gamma}\right]$ and all previous calculations are correct under this assumption and the solution to the HJBE must be the value function. Of course one can still check the admissibility of optimal control and that the stochastic integral is actually a MG but since it's trivial we neglect the procedure.

Example: Production-Consumption Model

Consider the infinite horizon model for a firm with capital value K_t, invvestment rate I_t and stock price S_t per unit of capital at time t. Assume that within the time interval $[t, t+h]$ $(h \to 0^+)$ the firm maintains its investment rate I_t, so

$$K_{t+h} = K_t \frac{S_{t+h}}{S_t} + hI_t \tag{389}$$

this gives the following SDE that

$$dK_t = K_t \frac{dS_t}{S_t} + I_t\, dt \tag{390}$$

Now the debt amount L_t of the firm is affected by the interest rate r, the consumption rate C_t and the productivity rate P_t of capital. Consider an infinitesimal time increment $h \to 0^+$ to get

$$L_{t+h} = L_t \cdot e^{rh} + hC_t + hI_t - (P_{t+h} - P_t)\frac{K_t}{S_t} \tag{391}$$

described by the following SDE that

$$dL_t = rL_t\, dt + (C_t + I_t)\, dt - \frac{K_t}{S_t}\, dP_t \tag{392}$$

The dynamics of stock price S_t and productivity rate of capital P_t are known by

$$\frac{dS_t}{S_t} = \mu\, dt + \sigma_1\, dB_t^1 \tag{393}$$

$$dP_t = b\, dt + \sigma_2\, dB_t^2 \tag{394}$$

where $B_t = (B_t^1, B_t^2)$ is 2-dimensional BM. For the convenience of notation, set $Y_t = \log S_t$ and apply Ito formula to see

$$dY_t = \left(\mu - \frac{\sigma_1^2}{2}\right) dt + \sigma_1\, dB_t^1 \tag{395}$$

we will use the dynamic of Y_t instead of S_t in the following context. The firm has net value

$$X_t = K_t - L_t \tag{396}$$

and we want to select **the control variables** $k_t = \frac{K_t}{X_t}, c_t = \frac{C_t}{X_t}$, the percentage of net value to invest and consume such that the net value of the company is maximized.

Since X_t, Y_t are observable states, we want to get the SDE system that describes the evolution of the states under the control variables

$$\begin{cases} dX_t = X_t[k_t(\mu - r + be^{-Y_t}) + r - c_t]\,dt + \sigma_1 k_t X_t\,dB_t^1 + \sigma_2 k_t X_t e^{-Y_t}\,dB_t^2 \\ dY_t = \left(\mu - \frac{\sigma_1^2}{2}\right)\,dt + \sigma_1\,dB_t^1 \end{cases} \tag{397}$$

note that this system contains only process X_t, Y_t except the control variables. Note that we can replace L_t by $L_t = K_t - X_t = (k_t - 1)X_t$. For infinite horizon problems, we always need a discount factor $\beta > 0$ and we still select the CRRA utility function $U(x) = \frac{x^{1-\gamma}}{1-\gamma}$ $\gamma \in (0,1)$ that turns the consumption rate $C_t = c_t X_t$ into the utility. So now the expected utility following control k, c on observing state x, y is

$$J(x,y,k,c) = \mathbb{E}\left[\int_0^\infty e^{-\beta t} U(c_t X_t^{0,x,y,k,c})\,dt\right] \tag{398}$$

where $X_t^{0,x,y,k,c}$ denotes process X_t as the solution to the SDE system above with initial value condition $X_0 = x, Y_0 = y$. Note that **the value of Y_t affects the dynamics of** X_t, that's why we have to include y in the superscript. On the other hand, since the dynamics of Y_t does not depend on X_t, $Y_t^{0,y,k,c}$ is enough for notation purpose.

For the admissible set of control, we still accept the Markovian setting with integrability conditions

$$\forall (k,c) \in \mathscr{A}, \forall T > 0, \int_0^T k_t^2 + c_t^2\,dt < \infty\ a.s., \mathbb{E}\left[\int_0^\infty e^{-\beta t} U(c_t X_t^{0,x,y,k,c})\,dt\right] < \infty, k_t, c_t \geq 0 \tag{399}$$

so the value function is formed as

$$v(x,y) = \sup_{(k,c)\in\mathscr{A}} \mathbb{E}\left[\int_0^\infty e^{-\beta t} U(c_t X_t^{0,x,y,k,c})\,dt\right] \tag{400}$$

Let's then write out the HJBE of the value function

$$-\beta v(x,y) + \sup_{(k,c)\in\mathscr{A}} \left\{L^{(k,c)} v(x,y) + U(cx)\right\} = 0 \tag{401}$$

with infinitesimal generator as

$$L^{(k,c)} v(x,y) = C_{Drift} \cdot \nabla v + \frac{1}{2} tr(C_{Diff} C_{Diff}^T \nabla^2 v) \tag{402}$$

where C_{Drift} is the drift coefficient and C_{Diff} is the diffusion coefficient. In this case, $C_{Drift} \in \mathbb{R}^2, C_{Diff} \in \mathbb{R}^{2\times 2}$ and

$$C_{Drift} = \begin{bmatrix} x[k(\mu - r + be^{-y}) + r - c] \\ \mu - \frac{\sigma_1^2}{2} \end{bmatrix} \tag{403}$$

$$C_{Diff} = \begin{bmatrix} \sigma_1 kx & \sigma_2 kxe^{-y} \\ \sigma_1 & 0 \end{bmatrix} \tag{404}$$

so by doing some calculations, we figure out that

$$L^{(k,c)}v(x,y) = x[k(\mu - r + be^{-y}) + r - c] \cdot \partial_x v(x,y) + \left(\mu - \frac{\sigma_1^2}{2}\right) \cdot \partial_y v(x,y) \tag{405}$$

$$+ \frac{\sigma_1^2 k^2 x^2 + \sigma_2^2 k^2 x^2 e^{-2y}}{2} \cdot \partial_{xx} v(x,y) + \sigma_1^2 kx \cdot \partial_{xy} v(x,y) + \frac{\sigma_1^2}{2} \cdot \partial_{yy} v(x,y) \tag{406}$$

so the HJBE is written in the simpler form with the sup w.r.t. k and c torn apart

$$\beta \cdot v - \left(\mu - \frac{\sigma_1^2}{2}\right) \cdot \partial_y v - rx \cdot \partial_x v - \frac{\sigma_1^2}{2} \cdot \partial_{yy} v - \sup_{c \geq 0} \{U(cx) - cx \cdot \partial_x v\} \tag{407}$$

$$- \sup_{k \geq 0} \left\{ xk(\mu - r + be^{-y}) \cdot \partial_x v + \frac{\sigma_1^2 k^2 x^2 + \sigma_2^2 k^2 x^2 e^{-2y}}{2} \cdot \partial_{xx} v + \sigma_1^2 kx \cdot \partial_{xy} v \right\} = 0 \tag{408}$$

Note that the SDE for X_t has the form that $\frac{dX_t}{X_t}$ does not contain X_t any longer, so if Y_t is known this is just a Black-Scholes model and we would expect to see the solution as $X_t^{0,x,y,k,c} = x \cdot e^{Z(Y_t^{0,y,k,c},k,c)}$ for some function Z. That's why we say that $v(x,y) = x^{1-\gamma} v(1,y)$ from the expression of the value function and the CRRA utility function. As a result, we try the **ansatz** $v(x,y) = \frac{x^{1-\gamma}}{1-\gamma} e^{\phi(y)}$. Now let's first get rid of the sup in the HJBE

$$\max_c U(cx) - cx \cdot \partial_x v \tag{409}$$

$$c^* = (\partial_x v)^{-\frac{1}{\gamma}} \cdot \frac{1}{x} \tag{410}$$

$$\max_k xk(\mu - r + be^{-y}) \cdot \partial_x v + \frac{\sigma_1^2 k^2 x^2 + \sigma_2^2 k^2 x^2 e^{-2y}}{2} \cdot \partial_{xx} v + \sigma_1^2 kx \cdot \partial_{xy} v \tag{411}$$

$$k^* = \max\left\{0, -\frac{\sigma_1^2 x \cdot \partial_{xy} v + x(\mu - r + be^{-y}) \cdot \partial_x v}{(\sigma_1^2 x^2 + \sigma_2^2 x^2 e^{-2y}) \cdot \partial_{xx} v}\right\} \tag{412}$$

and plug in our ansatz to know

$$\begin{cases} c^* = e^{-\frac{\phi(y)}{\gamma}} \\ k^* = \max\left\{0, \frac{\sigma_1^2 \phi'(y) + \mu - r + be^{-y}}{\gamma(\sigma_1^2 + \sigma_2^2 e^{-2y})}\right\} \end{cases} \tag{413}$$

the next step is to solve $\phi(y)$. HJBE now becomes

$$\beta - \left(\mu - \frac{\sigma_1^2}{2}\right) \phi'(y) - r(1-\gamma) - \frac{\sigma_1^2}{2}[(\phi'(y))^2 + \phi''(y)] - \sup_{c \geq 0} \left\{c^{1-\gamma} e^{-\phi(y)} - c(1-\gamma)\right\} \tag{414}$$

$$-(1-\gamma) \sup_{k \geq 0} \left\{ k(\mu - r + be^{-y} + \sigma_1^2 \phi'(y)) - \frac{\gamma k^2}{2}(\sigma_1^2 + e^{-2y} \sigma_2^2) \right\} = 0 \tag{415}$$

it's hard to solve this ODE for $\phi(y)$, but we can argue that this ODE has **unique C^2 bounded solution $\phi(y)$ for large enough** β and the following analysis will be conducted without knowing the form of $\phi(y)$.

Verification of Production-Consumption Model

Now let's prove that verification theorem holds for this problem so the solution to HJBE is exactly the value function. However, the difficulty is that we have to **conduct analysis without knowing the form of** $\phi(y)$. Firstly, the solution $v(x,y) = \frac{x^{1-\gamma}}{1-\gamma}e^{\phi(y)}$ is C^2 since $\phi \in C^2$. We have derived the optimal control c^*, k^* and they are well-defined functions for each fixed pair (x,y) and the condition $\overline{\lim}_{T\to\infty} e^{-\beta T}\cdot \mathbb{E}v(X_T^{0,x,y,k,c}, Y_T^{0,y,k,c}) \geq 0$ is naturally satisfied due to non-negativity of v on $\mathbb{R}^+ \times \mathbb{R}$. Now it's also obvious that the growth condition is satisfies since v is of order $x^{1-\gamma}$ in x and bounded in y. The only conditions to verify are $\underline{\lim}_{T\to\infty} e^{-\beta T}\cdot \mathbb{E}v(X_T^{0,x,y,k^*,c^*}, Y_T^{0,y,k^*,c^*}) \leq 0$ and the admissibility.

Note that the optimal control is given by

$$\begin{cases} c_t^* = e^{-\frac{\phi(Y_t)}{\gamma}} \\ k_t^* = \max\left\{0, \frac{\sigma_1^2\phi'(Y_t)+\mu-r+be^{-Y_t}}{\gamma(\sigma_1^2+\sigma_2^2e^{-2Y_t})}\right\} \end{cases} \tag{416}$$

only has something to do with the observation of Y_t, so it satisfies the **measurability** condition and

$$\forall T>0, \int_0^T (k_t^*)^2 + (c_t^*)^2\, dt < \infty \ a.s. \tag{417}$$

on the other hand, notice that if y is the critical point of ϕ', then $\phi''(y) = 0$ so

$$\beta - \left(\mu - \frac{\sigma_1^2}{2}\right)\phi'(y) - r(1-\gamma) - \frac{\sigma_1^2}{2}[\phi'(y)]^2 - \sup_{c\geq 0}\left\{c^{1-\gamma}e^{-\phi(y)} - c(1-\gamma)\right\} \tag{418}$$

$$-(1-\gamma)\sup_{k\geq 0}\left\{k(\mu - r + be^{-y} + \sigma_1^2\phi'(y)) - \frac{\gamma k^2}{2}(\sigma_1^2 + e^{-2y}\sigma_2^2)\right\} = 0 \tag{419}$$

$$\beta - \left(\mu - \frac{\sigma_1^2}{2}\right)\phi'(y) - r(1-\gamma) - \frac{\sigma_1^2}{2}[\phi'(y)]^2 - \sup_{c\geq 0}\left\{c^{1-\gamma}e^{-\phi(y)} - c(1-\gamma)\right\} \geq 0 \tag{420}$$

$$\beta - \left(\mu - \frac{\sigma_1^2}{2}\right)\phi'(y) - r(1-\gamma) - \frac{\sigma_1^2}{2}[\phi'(y)]^2 - (1-\gamma)e^{\frac{\gamma}{\gamma-1}\phi(y)} \geq 0 \tag{421}$$

we have proved that ϕ' is also bounded, so now $c_t^*, k_t^*, k_t^*e^{-Y_t}$ are all bounded. Let's look back at the dynamics of X_t that

$$\frac{dX_t}{X_t} = [k_t(\mu - r + be^{-Y_t}) + r - c_t]\, dt + \sigma_1 k_t\, dB_t^1 + \sigma_2 k_t e^{-Y_t}\, dB_t^2 \tag{422}$$

all coefficients are bounded, so the expectation of X_t given optimal control must be dominated by the expectation of some geometric BM (if all coefficients are constant, this is just BS model), i.e.

$$\exists M>0, \forall t>0, \mathbb{E}(X_t^{0,x,y,k^*,c^*})^2 \leq x^2\cdot e^{Mt} \tag{423}$$

the result is that

$$\mathbb{E}\left[\int_0^\infty e^{-\beta t}U(c_tX_t^{0,x,y,k^*,c^*})\,dt\right] \le C\cdot\int_0^\infty e^{-\beta t}\cdot\mathbb{E}\left[(X_t^{0,x,y,k^*,c^*})^2\right]\,dt \tag{424}$$

$$\le Cx^2\cdot\int_0^\infty e^{(M-\beta)t}\,dt < \infty\ (if\ \beta > M) \tag{425}$$

we have proved that the optimal control also satisfies the **integrability** conditions, so k^*, c^* are admissible optimal controls for large enough β.

Let's consider the last condition

$$\underline{\lim}_{T\to\infty}e^{-\beta T}\cdot\mathbb{E}v(X_T^{0,x,y,k^*,c^*}, Y_T^{0,y,k^*,c^*}) \le C\cdot\underline{\lim}_{T\to\infty}e^{-\beta T}\cdot\mathbb{E}[X_T^{0,x,y,k^*,c^*}]^{1-\gamma} \tag{426}$$

$$\le C\cdot\underline{\lim}_{T\to\infty}e^{-\beta T}\cdot\mathbb{E}[X_T^{0,x,y,k^*,c^*}]^2 \tag{427}$$

$$\le Cx^2\cdot\underline{\lim}_{T\to\infty}e^{(M-\beta)T} \tag{428}$$

$$= 0\ (if\ \beta > M) \tag{429}$$

To conclude, **when β is large enough** (larger than a fixed constant M that appears in the exponential of the expectation of a geometric BM that dominates $\mathbb{E}(X_t^{0,x,y,k^*,c^*})^2)$, the verification theorem holds and the production-consumption model is solved, with $c_t^* = c_t^*(Y_t), k_t^* = k_t^*(Y_t)$ as optimal Markovian controls.

Remark. *In this example, we will have to put up an ansatz, use ODE techniques to prove the property of the solution without solving the ODE, and prove that the conditions of verification theorem hold with some tricks. This is the general frame of the PDE approach to stochastic control problem, where we cannot solve out the optimal control and the HJBE directly.*

Backward Stochastic Differential Equation (BSDE)

In order to introduce the BSDE approach to stochastic control problems, let's first start by introducing some notations and restating the existence and uniqueness of the strong solution to SDE.

Let $\mathbb{H}^{0,k}$ denote the collection of all progressively measurable process that takes values in $\mathbb{R}^k$, the Hilbert space the solutions are living in is denoted as

$$\mathbb{H}^{2,k} = \left\{ Z \in \mathbb{H}^{0,k} : \mathbb{E}\int_0^T ||Z_s||^2 \, ds < \infty \right\} \tag{430}$$

where **the time horizon is always finite in $[0,T]$ for a large enough fixed constant T**. Our interest lies in the strong solution to the SDE that looks like

$$dX_t = b(t, X_t)\, dt + \sigma(t, X_t)\, dB_t \tag{431}$$

where $X_t, b(t, X_t)$ takes values in $\mathbb{R}^d$ and $\sigma(t, X_t)$ takes values in $\mathbb{R}^{d\times m}$, with $B_t = (B_t^1, ..., B_t^m)$ as an m-dimensional BM. By mentioning the strong solution, we **define that X_t is a strong solution if** (i): it satisfies the SDE above (ii): it also satisfies the **integrability condition** that

$$\int_0^T ||b(t, X_t)|| + ||\sigma(t, X_t)||^2 \, dt < \infty \ a.s. \tag{432}$$

Existence and Uniqueness of Strong Solution

In the stochastic calculus notes we have proved the existence and uniqueness of the strong solution to such SDE under growth condition and Lipschitz condition with the Picard iteration technique. However, here we apply some different assumptions and also derive some moment estimates on the solution with some difference techniques. In the following context, **it's always assumed that** $\forall x \in \mathbb{R}^d, \{b(t,x)\}_{t\in[0,T]} \in L^2([0,T]), \{\sigma(t,x)\}_{t\in[0,T]} \in L^2([0,T])$ **and b, σ are both Lipschitz in variable x.**

Theorem 17. ***(Existence and Uniqueness of Strong Solution)*** *Assume that $X_0 \in L^2$ is independent of the BM B_t and b, σ satisfies the assumptions above, then there exists a unique solution of SDE in $\mathbb{H}^{2,d}$ such that for some constant $C = C(T, Lips(b), Lips(\sigma)) > 0$,*

$$\mathbb{E} \sup_{t\in[0,T]} ||X_t||^2 \leq C(1 + \mathbb{E}||X_0||^2)e^{CT} \tag{433}$$

Proof. Instead of constructing the Picard iteration sequence, let's form the solution to the SDE as a fixed point and apply the contraction mapping theorem. Set the mapping $U : \mathbb{H}^{2,d} \to \mathbb{H}^{2,d}$ as

$$U(X)_t = X_0 + \int_0^t b(s, X_s)\, ds + \int_0^t \sigma(s, X_s)\, dB_s \tag{434}$$

so X_t is the solution to the SDE if and only if $X = U(X)$, i.e. X is the fixed point of U. Now we only have to prove that U is a contraction mapping under the Hilbert space norm induced by the inner product $\langle X, Y\rangle_{\mathbb{H}^{2,d}} = \mathbb{E}\int_0^T X_t \cdot Y_t \, dt$ $(\forall X, Y \in \mathbb{H}^{2,d})$.

However, the first problem we have to face is that we don't know if $U(X)$ is necessarily in $\mathbb{H}^{2,d}$, so we have to estimate $||U(X)||^2$.

$$||U(X)||^2 = \mathbb{E}\int_0^T ||U(X)_t||^2 \, dt \tag{435}$$

$$\leq 3\int_0^T \mathbb{E}||X_0||^2 \, dt + 3\mathbb{E}\int_0^T ||\int_0^t b(s, X_s)\, ds||^2 \, dt + 3\mathbb{E}\int_0^T ||\int_0^t \sigma(s, X_s)\, dB_s||^2 \, dt \tag{436}$$

the first term is obviously finite since $X_0 \in L^2$. To estimate the other two terms, we have to use the Lipschitz property in the way that $\Big| ||b(t,x)|| - ||b(t,0)|| \Big| \leq ||b(t,x) - b(t,0)|| \leq Lips(b) \cdot ||x||$ so $||b(t,x)||^2 \leq C(1 + ||b(t,0)||^2 + ||x||^2)$ for some constant $C = C(T, Lips(b), Lips(\sigma)) > 0$ and the same argument holds for σ. That's why we have

$$\mathbb{E}\int_0^T ||\int_0^t b(s, X_s)\, ds||^2 \, dt \leq \mathbb{E}\int_0^T t \cdot \int_0^t ||b(s, X_s)||^2 \, ds \, dt \tag{437}$$

$$\leq CT \cdot \mathbb{E}\int_0^T \int_0^t (1 + ||b(s,0)||^2 + ||X_s||^2)\, ds \, dt \tag{438}$$

$$\leq CT^2 \cdot \int_0^T (1 + ||b(s,0)||^2 + \mathbb{E}||X_s||^2)\, ds \tag{439}$$

$$\leq CT^3 \cdot \left(1 + ||b(\cdot, 0)||^2 + \sup_{s\in[0,T]} \mathbb{E}||X_s||^2\right) \tag{440}$$

by Cauchy inequality and $||b(\cdot, 0)||^2 < \infty$ since $\forall x \in \mathbb{R}^d, \{b(t,x)\}_{t\in[0,T]} \in L^2([0,T])$. Similarly, for the third term,

$$\mathbb{E}\int_0^T ||\int_0^t \sigma(s, X_s)\, dB_s||^2 \, dt \leq T \cdot \mathbb{E}\sup_{t\in[0,T]} ||\int_0^t \sigma(s, X_s)\, dB_s||^2 \tag{441}$$

$$\leq 4T \cdot \mathbb{E}||\int_0^T \sigma(s, X_s)\, dB_s||^2 \tag{442}$$

$$\leq 4T \cdot \mathbb{E}\int_0^T ||\sigma(s, X_s)||^2 \, ds \tag{443}$$

$$\leq 4CT^2 \cdot \left(1 + ||\sigma(\cdot, 0)||^2 + \sup_{s\in[0,T]} \mathbb{E}||X_s||^2\right) \tag{444}$$

by Doob's L^p inequality. Since $X \in \mathbb{H}^{2,d}$, $\sup_{s\in[0,T]} \mathbb{E}||X_s||^2 < \infty$ and it's proved that the image of U is still in $\mathbb{H}^{2,d}$.

To prove the contraction mapping property, let's first introduce another norm on $\mathbb{H}^{2,d}$ as

$$||X||_\alpha^2 = \mathbb{E}\int_0^T e^{-\alpha t}||X_t||^2 \, dt \ (\alpha > 0) \tag{445}$$

which is actually equivalent to the original norm (when $\alpha = 0$). Under this norm, consider

$$
\begin{aligned}
||U(X) - U(Y)||_\alpha^2 &= \mathbb{E}\int_0^T e^{-\alpha t}||U(X)_t - U(Y)_t||^2\,dt &(446)\\
&= \int_0^T e^{-\alpha t}\mathbb{E}||U(X)_t - U(Y)_t||^2\,dt &(447)\\
&\le 2\int_0^T e^{-\alpha t}\cdot\left[\mathbb{E}||\int_0^t b(s,X_s) - b(s,Y_s)\,ds||^2 + \mathbb{E}||\int_0^t \sigma(s,X_s) - \sigma(s,Y_s)\,dB_s||^2\right]dt &(448)\\
&\le 2\int_0^T e^{-\alpha t}\cdot\left[t\cdot\mathbb{E}\int_0^t ||b(s,X_s) - b(s,Y_s)||^2\,ds + \mathbb{E}\int_0^t ||\sigma(s,X_s) - \sigma(s,Y_s)||^2\,ds\right]dt &(449)\\
&\le C\int_0^T e^{-\alpha t}\cdot\int_0^t \mathbb{E}||X_s - Y_s||^2\,ds\,dt &(450)\\
&= C\int_0^T \mathbb{E}||X_s - Y_s||^2\cdot\int_s^T e^{-\alpha t}\,dt\,ds &(451)\\
&\le \frac{C}{\alpha}||X - Y||_\alpha^2 &(452)
\end{aligned}
$$

so if α is large enough, $\frac{C}{\alpha} < 1$ for some constant C that only depends on $T, Lips(b), Lips(\sigma)$, the mapping U is a strict contraction mapping (since for any $\alpha \ge 0$, the norms are equivalent). This proved the existence and the uniqueness of the solution immediately by contraction mapping theorem.

For the moment estimate of the solution,

$$
\begin{aligned}
\forall t\in[0,T], \mathbb{E}\sup_{s\in[0,t]}||X_s||^2 &= \mathbb{E}\sup_{s\in[0,t]}||X_0 + \int_0^s b(r,X_r)\,dr + \int_0^s \sigma(r,X_r)\,dB_r||^2 &(453)\\
&\le 3\mathbb{E}||X_0||^2 + 3\mathbb{E}\sup_{s\in[0,t]}||\int_0^s b(r,X_r)\,dr||^2 + 3\mathbb{E}\sup_{s\in[0,t]}||\int_0^s \sigma(r,X_r)\,dB_r||^2 &(454)\\
&\le 3\mathbb{E}||X_0||^2 + 3t\cdot\mathbb{E}\int_0^t ||b(r,X_r)||^2\,dr + 12\int_0^t \mathbb{E}||\sigma(r,X_r)||^2\,dr &(455)\\
&\le C\cdot\left(1 + \mathbb{E}||X_0||^2 + \int_0^t \mathbb{E}\sup_{p\in[0,r]}||X_p||^2\,dr\right) &(456)
\end{aligned}
$$

with Doob's L^p inequality applied once more. Now notice that by denoting $f(t) = \mathbb{E}\sup_{s\in[0,t]}||X_s||^2$, we have that $f(t) \le C\cdot\left(1 + \mathbb{E}||X_0||^2 + \int_0^t f(r)\,dr\right)$, Grownwall's inequality proves the conclusion that

$$
\forall t\in[0,T], \mathbb{E}\sup_{s\in[0,t]}||X_s||^2 \le C(1 + \mathbb{E}||X_0||^2)e^{Ct} \tag{457}
$$

□

Remark. *From the uniqueness argument in the contraction mapping theorem, if two processes X_t, Y_t are both solutions to the same SDE, then $||X_t - Y_t||_{\mathbb{H}^{2,d}} = 0$, so $X_t = Y_t$ a.a.$(t,\omega)\in[0,\infty)\times\Omega$, this is a little bit weaker than* ***the uniqueness in the sense of modification****.*

Remark. *The norm equivalency argument is a useful trick for finite time-horizon processes. It's not hard to verify that* $||\cdot||_\alpha$ *is actually a norm on* $\mathbb{H}^{2,d}$ *and that*

$$e^{-\alpha T}||X||_{\mathbb{H}^{2,d}}^2 \le ||X||_\alpha^2 \le ||X||_{\mathbb{H}^{2,d}}^2 \tag{458}$$

however, the freedom in choosing α *ensures that there exists a contraction mapping with the contraction coefficient* $\frac{C}{\alpha} < 1$.

Settings of BSDE

BSDE is a special kind of SDE for which the terminal condition is given as a certain random variable so the evolution shall actually proceed backwardly. The BSDE consists of a **driver (coefficient)** $\Psi : [0,T]\times\Omega\times\mathbb{R}^p\times\mathbb{R}^{pm} \to \mathbb{R}^p$ such that $\forall (y,z) \in \mathbb{R}^p \times \mathbb{R}^{pm}$, $\{\Psi(t,y,z)\}_{t\in[0,T]}$ is $\mathscr{P}$ measurable for the sigma field $\mathscr{P}$ on $[0,T]\times\Omega$ generated by $\{\mathscr{F}_t\}_{t\in[0,T]}$ measurable bounded process. The filtration $\{\mathscr{F}_t\}_{t\in[0,T]}$ is that of the m-dimensional BM B_t by default. The BSDE also has a known **terminal condition** $\xi \in L^2$. The BSDE for Y_t, Z_t is always formed as

$$\begin{cases} -dY_t = \Psi(t, Y_t, Z_t)\, dt - Z_t\, dB_t \\ Y_T = \xi \end{cases} \tag{459}$$

and is denoted $BSDE(\Psi, \xi)$.

The **solution** to such BSDE is defined as a pair of process (Y, Z) such that it satisfies the BSDE

$$Y_t = \xi + \int_t^T \Psi(s, Y_s, Z_s)\, ds - \int_t^T Z_s\, dB_s \tag{460}$$

with some regularity conditions that

$$Y \in \mathbb{S}^{2,p}, Z \in \mathbb{H}^{2,pm} \tag{461}$$

where $\mathbb{H}^{2,pm}$ stands for the Hilbert space of process taking values in $\mathbb{R}^{pm}$ defined above and

$$\mathbb{S}^{2,p} = \left\{ Y \in \mathbb{H}^{0,p} : \mathbb{E} \sup_{t\in[0,T]} ||Y_t||^2 < \infty \right\} \tag{462}$$

Remark. *Let's briefly talk about the* ***motivation*** *of BSDE. One might be confused with the reason why there is a pair of processes appearing in the BSDE instead of a single process which is the case for forward SDE. Actually, the motivation comes from the generalization of the Feynman-Kac formula.*

Recall that the Feynman-Kac formula provides the probabilistic characterization of the solution to the PDE. However, the most general case to deal with is the ***linear parabolic PDE*** *that looks like*

$$\partial_t u + Lu + fu + g = 0 \tag{463}$$

for some nice enough f, g, the infinitesimal generator L and a given terminal condition (note that Feynman-Kac formula deals with a PDE with given terminal condition).

Now we hope to find the probabilistic characterization for the solution to ***semilinear parabolic PDE*** *that looks like*

$$\partial_t u + Lu + \Psi(t, x, u, \partial_x u) = 0 \tag{464}$$

and we hope to connect the solution u (assume it exists) with the Ito diffusion X_t generated by the dynamics

$$dX_t = b(X_t)\,dt + \sigma(X_t)\,dB_t \tag{465}$$

whose infinitesimal generator is exactly L and has a given initial value condition. Recall how we connect solution to PDE with diffusion, we shall perturb the initial value condition of the diffusion and derive a PDE. That's why we first assume that $X^{t,x}$ denotes the Ito diffusion generated by given initial condition $X_t = x$ and $Y_s^{t,x} = u(s, X_s^{t,x})$. By Ito formula,

$$u(t+h, X_{t+h}^{t,x}) = u(t,x) + \int_t^{t+h} (\partial_t + L)u(s, X_s^{t,x})\,ds + \int_t^{t+h} \partial_x u(s, X_s^{t,x})\sigma(X_s^{t,x})\,dB_s \tag{466}$$

so the dynamics of Y is

$$\begin{aligned} dY_t &= (\partial_t + L)u(t, X_t)\,dt + \partial_x u(t, X_t)\sigma(X_t)\,dB_t &(467)\\ &= -\Psi(t, X_t, u(t, X_t), \partial_x u(t, X_t))\,dt + \partial_x u(t, X_t)\sigma(X_t)\,dB_t &(468)\end{aligned}$$

As a result, if we slightly modify the components of Ψ into $\Psi(t, X_t, u(t, X_t), \partial_x u(t, X_t)\sigma(X_t))$, i.e. such $\Psi = \Psi(t, x, u, \partial_x u \cdot \sigma(x))$ depends on $\partial_x u$ in the way that it only depends on the product $\partial_x u \cdot \sigma(x)$. Then by setting

$$\begin{cases} Y_t = u(t, X_t) \\ Z_t = \partial_x u(t, X_t)\sigma(X_t) \end{cases} \tag{469}$$

we see the form of the BSDE that

$$dY_t = -\Psi(t, X_t, Y_t, Z_t)\,dt + Z_t\,dB_t \tag{470}$$

which explains why the BSDE we consider has such a form. We will see exactly this form again in the following nonlinear Feynman-Kac formula.

Existence and Uniqueness of the Solution to BSDE

Now let's **assume that** $\Psi(t, y, z)$ **is Lipschitz in** (y, z) **and** $\{\Psi(t, 0, 0)\}_{t\in[0,T]} \in \mathbb{H}^{2,p}$, and these assumptions suffice to ensure the existence and uniqueness of the solution to BSDE by a similar fixed point argument.

Theorem 18. ***(Existence and Uniqueness of Solution to BSDE)*** *Under the assumption above, the BSDE*

$$\begin{cases} -dY_t = \Psi(t, Y_t, Z_t)\,dt - Z_t\,dB_t \\ Y_T = \xi \end{cases} \tag{471}$$

has unique solution.

Proof. Define the operator $U : \mathbb{S}^{2,p} \times \mathbb{H}^{2,pm} \to \mathbb{S}^{2,p} \times \mathbb{H}^{2,pm}$ such that it maps (Y, Z) to $(\tilde{Y}, \tilde{Z})$ in a way that

$$\tilde{Y}_t = \xi + \int_t^T \Psi(s, Y_s, Z_s)\,ds - \int_t^T \tilde{Z}_s\,dB_s \tag{472}$$

note that this construction provides $\tilde{Y}, \tilde{Z}$ simultaneously. Consider the L^2 MG

$$M_t = \mathbb{E}\left[\xi + \int_0^T \Psi(s, Y_s, Z_s)\,ds \middle| \mathscr{F}_t\right] \tag{473}$$

and apply the MG representation theorem to find that there exists unique $\tilde{Z} \in \mathbb{H}^{2,pm}$ such that

$$M_t = M_0 + \int_0^t \tilde{Z}_s\,dB_s \tag{474}$$

as a result, set

$$\tilde{Y}_t = \mathbb{E}\left[\xi + \int_t^T \Psi(s, Y_s, Z_s)\,ds \middle| \mathscr{F}_t\right] = M_t - \int_0^t \Psi(s, Y_s, Z_s)\,ds \tag{475}$$

to find that $\tilde{Y}_T = \xi$ and

$$d\tilde{Y}_t = dM_t - \Psi(t, Y_t, Z_t)\,dt = \tilde{Z}_t\,dB_t - \Psi(t, Y_t, Z_t)\,dt \tag{476}$$

so the mapping U is well-defined. The fixed point of U is just the solution to the BSDE. Before verifying the contraction mapping property, let's first verify that the image of U is in the space $\mathbb{S}^{2,p} \times \mathbb{H}^{2,pm}$. Note that by MG representation theorem, we have already proved that $\tilde{Z} \in \mathbb{H}^{2,pm}$, so we shall estimate the moment of $\tilde{Y}$

$$\mathbb{E}\sup_{t\in[0,T]}||\tilde{Y}_t||^2 \le C\left[\mathbb{E}||\xi||^2 + \mathbb{E}\sup_{t\in[0,T]}||\int_t^T \Psi(s,Y_s,Z_s)\,ds||^2 + \mathbb{E}\sup_{t\in[0,T]}||\int_t^T \tilde{Z}_s\,dB_s||^2\right] \tag{477}$$

$$\le C\left[\mathbb{E}||\xi||^2 + T\cdot\mathbb{E}\int_0^T ||\Psi(s,Y_s,Z_s)||^2\,ds + \mathbb{E}||\int_0^T \tilde{Z}_s\,dB_s||^2 + \mathbb{E}\sup_{t\in[0,T]}||\int_0^t \tilde{Z}_s\,dB_s||^2\right] \tag{478}$$

$$\le C\left[\mathbb{E}||\xi||^2 + \mathbb{E}\int_0^T ||\Psi(s,0,0)||^2\,ds + \mathbb{E}\int_0^T ||Y_s||^2\,ds + \mathbb{E}\int_0^T ||Z_s||^2\,ds + \mathbb{E}\int_0^T ||\tilde{Z}_s||^2\,ds\right] \tag{479}$$

$$< \infty \tag{480}$$

by Doob's L^p inequality, Ito's isometry and the Lipschitz assumption for some constant $C = C(T, Lips(\Psi)) > 0$. As a result, $\tilde{Y} \in \mathbb{S}^{2,p}$ and we only need to verify the contraction mapping property.

Recall the trick applied in previous proof for forward SDE that we have introduced a family of equivalent norms with the freedom to choose the parameter in the norm in order to ensure that the contraction coefficient is always strictly less than 1. Apply the same trick here to define the norm

$$||(Y,Z)||_\alpha^2 = \mathbb{E}\int_0^T e^{\alpha t}\cdot(||Y_t||^2 + ||Z_t||^2)\,dt \tag{481}$$

and notice that this family of norm is equivalent for $\alpha \ge 0$ and when $\alpha = 0$ we get the canonical norm on the product of two Hilbert spaces $\mathbb{H}^{2,p}\times\mathbb{H}^{2,pm}$. Eventually we can see that $\forall (Y^1,Z^1),(Y^2,Z^2) \in \mathbb{H}^{2,p}\times\mathbb{H}^{2,pm}$

$$||U(Y^1,Z^1) - U(Y^2,Z^2)||_\alpha^2 \le \frac{2c^2(T+1)}{\alpha}||(Y^1,Z^1)-(Y^2,Z^2)||_\alpha^2 \tag{482}$$

so we can take α to be large enough such that $\frac{2c^2(T+1)}{\alpha} < 1$ and the theorem is proved (calculations are ommitted here). □

Remark. *The **Burkholder-Davis-Gundy (BDG) inequality** is an important tool in continuous-time stochastic analysis. It asserts that for $\forall p > 0$, any continuous local MG M with $M_0 = 0$ and any stopping time τ, there always exists $c = c(p), C = C(p)$ such that*

$$c\cdot\mathbb{E}\langle M,M\rangle_\tau^{\frac{p}{2}} \le \mathbb{E}\sup_{0\le t\le\tau}|M_t|^p \le C\cdot\mathbb{E}\langle M,M\rangle_\tau^{\frac{p}{2}} \tag{483}$$

it enables us to deal with the expectation of the sup of stochastic integral since the expectation of the quadratic variation of stochastic integral is always easy to deal with. One will have to use BDG inequality in the calculations above to get rid of a local MG after taking expectation.

Example: Trivial Driver

Let's consider the trivial case where $\Psi = 0$ so $dY_t = Z_t\,dB_t$. In other words, the stochastic integral of Z w.r.t. BM gives Y and now we are given the terminal value of Y. This makes us recall the MG representation theorem since Z very much looks like the process whose stochastic intergal generates the MG.

Consider the L^2 MG

$$M_t = \mathbb{E}(\xi|\mathscr{F}_t) \tag{484}$$

so MG representation theorem concludes that there exists unique $Z \in \mathbb{H}^{2,pm}$ such that $M_t = M_0 + \int_0^t Z_s\,dB_s = \mathbb{E}\xi + \int_0^t Z_s\,dB_s$. Now set

$$Y_t = M_t \tag{485}$$

to see that $Y_T = \xi, dY_t = dM_t = Z_t\,dB_t$, so (Y, Z) is the solution pair. The existence and uniqueness of the solution guarantees that this is the only solution to such BSDE.

Remark. *This BSDE $dY_t = Z_t\,dB_t$ makes no sense when an initial value condition is given, let's say, $Y_0 = 0$ since we immediately know that*

$$Y_t = \int_0^t Z_s\,dB_s \tag{486}$$

and there are infinitely many solutions. For each selection of Z there always exists a Y such that this SDE holds. This example illustrates the difference between forward SDE and BSDE. BSDE is always related to ***MG representation theorem*** *since the solution is a pair of processes and one of them often has to be fixed first by MG representation.*

Example: Linear BSDE

The majority of BSDEs cannot be solved. However, we can consider a special type of BSDE where **the driver is linear in** y, z. We write the linear BSDE in the following form

$$\begin{cases} -dY_t = (P_tY_t + Q_t \cdot Z_t + R_t)\,dt - Z_t\,dB_t \\ Y_T = \xi \end{cases} \tag{487}$$

assuming $p = 1$ so this BSDE is in 1-dimension with $P_t \in \mathbb{R}, Q_t \in \mathbb{R}^m$ bounded. It's easy to verify that such BSDE satisfies the conditions listed in the theorem above so it has unique solution.

Theorem 19. ***(Solution to Linear BSDE)*** *The solution to linear BSDE is given by*

$$\Gamma_tY_t = \mathbb{E}\left[\Gamma_T\xi + \int_t^T \Gamma_sR_s\,ds\middle|\mathscr{F}_t\right] \tag{488}$$

where Γ *is the* ***adjoint process*** *given by*

$$\begin{cases} d\Gamma_t = \Gamma_t(P_t\,dt + Q_t\,dB_t) \\ \Gamma_0 = 1 \end{cases} \tag{489}$$

Proof. Since we already have existence and uniqueness of the solution, just need to construct the appropriate Z_t and show that the Y_t given makes (Y, Z) a solution pair.

Notice that $\Gamma_T Y_T = \xi\Gamma_T$ so the terminal condition satisfies. By Ito formula, $-dY_t = (P_tY_t + Q_t \cdot Z_t + R_t)\,dt - Z_t\,dB_t$ holds if and only if

$$d(\Gamma_t Y_t) = \Gamma_t\,dY_t + Y_t\,d\Gamma_t + d\langle Y, \Gamma\rangle_t \tag{490}$$

$$= -\Gamma_t R_t\,dt + \Gamma_t(Z_t + Y_tQ_t)\,dB_t \tag{491}$$

$$\Gamma_t Y_t - Y_0 = -\int_0^t \Gamma_s R_s\,ds + \int_0^t \Gamma_s(Z_s + Y_sQ_s)\,dB_s \tag{492}$$

to notice that the stochastic integral is actually a MG we shall calculate

$$\mathbb{E}\left\langle \int_0^\cdot \Gamma_s(Z_s + Y_sQ_s)\,dB_s, \int_0^\cdot \Gamma_s(Z_s + Y_sQ_s)\,dB_s \right\rangle_T = \mathbb{E}\int_0^T \Gamma_s^2||Z_s + Y_sQ_s||^2\,ds < \infty \tag{493}$$

since Q is bounded, $\mathbb{E}\sup_{t\in[0,T]}\Gamma_t^2 < \infty$ (since P_t, Q_t are both bounded) and that $Y \in \mathbb{S}^{2,1}, Z \in \mathbb{H}^{2,m}$. By BDG inequality, this tells us that

$$\mathbb{E}\sup_{t\in[0,T]}\left|\int_0^t \Gamma_s(Z_s + Y_sQ_s)\,dB_s\right| \le C\cdot\mathbb{E}\left[\int_0^T \Gamma_s^2||Z_s + Y_sQ_s||^2\,ds\right]^{\frac{1}{2}} < \infty \tag{494}$$

for some constant independent of T. The stochastic integral is a U.I. MG (the sup of local MG is integrable, satisfies the dominated condition). As a result, U.I. MG has to be closed

$$\Gamma_t Y_t + \int_0^t \Gamma_s R_s\,ds = \mathbb{E}\left[\Gamma_T Y_T + \int_0^T \Gamma_s R_s\,ds\middle|\mathscr{F}_t\right] = \mathbb{E}\left[\Gamma_T\xi + \int_0^T \Gamma_s R_s\,ds\middle|\mathscr{F}_t\right] \tag{495}$$

To let the BSDE hold for such Y, we only need to choose an appropriate $Z \in \mathbb{H}^{2,m}$ such that

$$Y_0 + \int_0^t \Gamma_s(Z_s + Y_sQ_s)\,dB_s = \mathbb{E}\left[\Gamma_T\xi + \int_0^T \Gamma_s R_s\,ds\middle|\mathscr{F}_t\right] \tag{496}$$

whose existence is guaranteed by the MG representation theorem! $\square$

Comparison Principles

Although general BSDE is very hard to solve, a comparison principle enables us to describe the property of the solution to BSDE without explicitly solving it. The comparison principle focuses on the comparison of the driver and the terminal conditions with few restrictions to apply.

Theorem 20. ***(Comparison Principles of BSDE)*** *$(\Psi^1, \xi^1), (\Psi^2, \xi^2)$ are two sets of drivers and terminal conditions for a BSDE with $(Y^1, Z^1), (Y^2, Z^2)$ as corresponding solution pairs. If now $BSDE(\Psi^2, \xi^2)$ satisfies the existence and uniqueness condition for the solution and $\xi^1 \leq \xi^2$ a.s., $\Psi^1(t, Y_t^1, Z_t^1) \leq \Psi^2(t, Y_t^1, Z_t^1)$ a.a.(t, ω) and $\Psi^2(t, Y_t^1, Z_t^1) \in \mathbb{H}^{2,p}$, then*

$$a.s.\ \forall t \in [0, T], Y_t^1 \leq Y_t^2 \tag{497}$$

In particular, if $Y_0^2 \leq Y_0^1$, then a.s. $\forall t \in [0, T], Y_t^1 = Y_t^2$.

Proof. Only prove for $p = 1$. Consider the difference $(P, Q) = (Y^2 - Y^1, Z^2 - Z^1)$, it satisfies

$$\begin{cases} -dP_t = [-\Psi^1(t, Y_t^1, Z_t^1) + \Psi^2(t, Y_t^2, Z_t^2)]\,dt - Q_t\,dB_t \\ P_T = \xi^2 - \xi^1 \end{cases} \tag{498}$$

rewrite the BSDE as a linear BSDE for (P, Q)

$$\begin{aligned} -dP_t &= [-\Psi^1(t, Y_t^1, Z_t^1) + \Psi^2(t, Y_t^2, Z_t^2)]\,dt - Q_t\,dB_t & (499)\\ &= \Big([\Psi^2(t, Y_t^1, Z_t^1) - \Psi^1(t, Y_t^1, Z_t^1)] + [\Psi^2(y, Y_t^2, Z_t^2) - \Psi^2(y, Y_t^1, Z_t^2)] & (500)\\ &+ [\Psi^2(y, Y_t^1, Z_t^2) - \Psi^2(y, Y_t^1, Z_t^1)]\Big)\,dt - Q_t\,dB_t & (501)\\ &= \Big([\Psi^2(t, Y_t^1, Z_t^1) - \Psi^1(t, Y_t^1, Z_t^1)] + \frac{\Psi^2(y, Y_t^2, Z_t^2) - \Psi^2(y, Y_t^1, Z_t^2)}{Y_t^2 - Y_t^1}\mathbb{I}_{Y_t^2 \neq Y_t^1} P_t & (502)\\ &+ \frac{\Psi^2(y, Y_t^1, Z_t^2) - \Psi^2(y, Y_t^1, Z_t^1)}{Z_t^2 - Z_t^1}\mathbb{I}_{Z_t^2 \neq Z_t^1} Q_t\Big)\,dt - Q_t\,dB_t & (503) \end{aligned}$$

let's denote the coefficients as

$$\begin{cases} \alpha_t = \Psi^2(t, Y_t^1, Z_t^1) - \Psi^1(t, Y_t^1, Z_t^1) \\ \beta_t = \frac{\Psi^2(y, Y_t^2, Z_t^2) - \Psi^2(y, Y_t^1, Z_t^2)}{Y_t^2 - Y_t^1}\mathbb{I}_{Y_t^2 \neq Y_t^1} \\ \gamma_t = \frac{\Psi^2(y, Y_t^1, Z_t^2) - \Psi^2(y, Y_t^1, Z_t^1)}{Z_t^2 - Z_t^1}\mathbb{I}_{Z_t^2 \neq Z_t^1} \end{cases} \tag{504}$$

to find that β_t, γ_t are both bounded since $\Psi^2(t, y, z)$ is Lipschitz in (y, z) and that $\alpha_t \in \mathbb{H}^{2,1}$ so

$$\begin{cases} -dP_t = [\beta_t P_t + \gamma_t Q_t + \alpha_t]\,dt - Q_t\,dB_t \\ P_T = \xi^2 - \xi^1 \end{cases} \tag{505}$$

is now a linear BSDE. By the previous theorem, the unique solution should be given by

$$\Gamma_t P_t = \mathbb{E}\left[\Gamma_T(\xi^2 - \xi^1) + \int_t^T \Gamma_s \alpha_s \, ds \Big| \mathscr{F}_t\right] \tag{506}$$

where Γ_t is the adjoint process such that

$$\begin{cases} d\Gamma_t = \Gamma_t(\beta_t \, dt + \gamma_t \, dB_t) \\ \Gamma_0 = 1 \end{cases} \tag{507}$$

since β_t, γ_t are both bounded, Γ_t is strictly positive and the condition tells us that almost surely $\xi^2 - \xi^1, \alpha_t$ are both non-negative. As a result, we conclude that almost surely P_t has to be non-negative, which proves the theorem.

Moreover, when $Y_0^2 \leq Y_0^1$, $P_0 \leq 0$ so $P_0 = 0$, which tells us that

$$\mathbb{E}\left[\Gamma_T(\xi^2 - \xi^1) + \int_0^T \Gamma_s \alpha_s \, ds\right] = 0 \tag{508}$$

so $a.s.$ $\xi^2 - \xi^1 = 0, \forall t \in [0, T], \alpha_t = 0$ and $\forall t \in [0, T], P_t = 0, Y_t^1 = Y_t^2$. □

Remark. *Note that the comparison principle is just an application of the solution to linear BSDE. The interesting point is that there's* ***no requirement on the regularity of*** (Ψ^1, ξ^1). ***As a corollary, if*** $\mathbb{P}\left(\xi^1 < \xi^2\right) > 0$ ***or*** $\Psi^1 < \Psi^2$ ***on a positive measure set under*** $dt \times d\mathbb{P}$***, then*** $Y_0^1 < Y_0^2$ ***gives a strict inequality for the initial values of*** Y.

Example

By taking $\Psi^1 = 0, \xi^1 = 0$, it's obvious that the solution to $BSDE(\Psi^1, \xi^1)$ is trivial $Y_t^1 = 0, Z_t^1 = 0$.

Apply the comparison principle to find out that if $\xi^2 \geq \xi^1 = 0$ $a.s.$, $\Psi^2(t, 0, 0) \geq \Psi^1(t, 0, 0) = 0$ $a.a.(t, \omega)$, then

$$a.s. \ \forall t \in [0, T], Y_t^2 \geq Y_t^1 = 0 \tag{509}$$

telling us that **if a BSDE has terminal condition almost surely non-negative and driver non-negative for almost all time and at** $(y, z) = (0, 0)$**, then the solution must be non-negative**. In particular, if $\mathbb{P}(\xi > 0) > 0$ or $\Psi^2(t, 0, 0) > 0$ $a.a.(t, \omega)$ then $Y_0 > 0$.

Stochastic Control Problem, BSDE Approach

BSDE and Value Function

Let's first consider the connection between the solution to BSDE and the value function. In stochastic control problems, the dynamics will be determined by a control parameter α so let's first consider the optimization of a family of BSDE. In the following context, the sup and inf taken for random variables are naturally specified as **essential sup and inf** (ignoring the zero measure set) and **all BSDEs appearing satisfies the existence and uniqueness condition**. We will be considering the minimization problem and the dynamics is given by BSDE as

$$\begin{cases} -dY_t = f^\alpha(t, Y_t, Z_t)\,dt - Z_t\,dB_t \\ Y_T = \xi^\alpha \end{cases} \tag{510}$$

where both the driver and the terminal condition depend on the control α in the admissible set $\mathscr{A}$.

Theorem 21. ***(Optimization of a Family of BSDE)*** *Let $(f, \xi), (f^\alpha, \xi^\alpha)$ be a family of driver-terminal condition pairs for $\alpha \in \mathscr{A}$ and $(Y, Z), (Y^\alpha, Z^\alpha)$ be the corresponding solution pairs. Suppose that $\exists \hat{\alpha} \in \mathscr{A}$ such that*

$$f(t, Y_t, Z_t) = \inf_\alpha f^\alpha(t, Y_t, Z_t) = f^{\hat{\alpha}}(t, Y_t, Z_t) \ a.a.(t, \omega) \tag{511}$$

$$\xi = \inf_\alpha \xi^\alpha = \xi^{\hat{\alpha}} \ a.s. \tag{512}$$

then

$$a.s.\ \forall t \in [0, T], Y_t = \inf_\alpha Y_t^\alpha = Y_t^{\hat{\alpha}} \tag{513}$$

Proof. By the comparison principles, for $\forall \alpha \in \mathscr{A}$, $\xi \leq \xi^\alpha, f(t, Y_t, Z_t) \leq f^\alpha(t, Y_t, Z_t)$ so the solutions to the BSDEs have the relationship that $a.s.\ \forall t \in [0, T], Y_t \leq Y_t^\alpha$. As a result, $a.s.\ \forall t \in [0, T], Y_t \leq \inf_\alpha Y_t^\alpha$.

By the existence of $\hat{\alpha} \in \mathscr{A}$, consider $(f^{\hat{\alpha}}, \xi^{\hat{\alpha}})$ to find that $(Y, Z), (Y^{\hat{\alpha}}, Z^{\hat{\alpha}})$ are both solution pairs to the BSDE with this driver and terminal condition. By uniqueness, $a.s.\ \forall t \in [0, T], Y_t = Y_t^{\hat{\alpha}}$ so $Y_t = Y_t^{\hat{\alpha}} \geq \inf_\alpha Y_t^\alpha$. $\square$

Remark. *It's easy to connect such Y_t^α to the problem value under control α. Then, $\hat{\alpha}$ stands for the optimal control and $Y_t^{\hat{\alpha}}$ is just the value function. The theorem above is telling us how to characterize the optimal control and the value function with BSDE, simply taking inf w.r.t. control for driver and terminal condition respectively will work.*

To see the correspondence between Y as the solution to BSDE and its structure as a value function, let's first set up the stochastic control problem. Assume that **the driver f is concave** and

$$F(t, b, c) = \sup_{(y,z)} \{f(t, y, z) - yb - zc\} \ \ (b, c) \in \mathbb{R} \times \mathbb{R}^m \tag{514}$$

as the Frenchel conjugate, so F is convex. The control variables are $\beta_t \in \mathbb{R}, \gamma_t \in \mathbb{R}^m$ taking values in the admissible set

$$\mathscr{A} = \left\{ (\beta, \gamma) : \mathbb{E} \int_0^T ||F(t, \beta_t, \gamma_t)||^2 \, dt < \infty, (\beta, \gamma) \text{ bounded progressive} \right\} \tag{515}$$

Since f is concave, we know that

$$f(t, y, z) = \inf_{(b,c)} \{F(t, b, c) + yb + zc\} \tag{516}$$

by the property of Frenchel conjugate. Naturally, we consider the family of linear drivers

$$f^{\beta,\gamma}(t, y, z) = F(t, \beta_t, \gamma_t) + y\beta_t + z\gamma_t \; (\beta, \gamma) \in \mathscr{A} \tag{517}$$

and assume that the solution to $BSDE(f^{\beta,\gamma}, \xi)$ is denoted $(Y^{\beta,\gamma}, Z^{\beta,\gamma})$

Theorem 22. ***(Solution to BSDE with Concave Driver as Value Function)*** *Let (Y, Z) denote the solution to $BSDE(f, \xi)$, then a.s. $\forall t \in [0, T], Y_t = \inf_{(\beta,\gamma)\in\mathscr{A}} Y_t^{\beta,\gamma}$ is the value function of such stochastic control problem and*

$$Y_t^{\beta,\gamma} = \mathbb{E}_{\mathbb{Q}^\gamma} \left[\int_t^T e^{\int_t^s \beta_u \, du} F(s, \beta_s, \gamma_s) \, ds + e^{\int_t^T \beta_u \, du} \xi \Big| \mathscr{F}_t \right] \tag{518}$$

where $\mathbb{Q}^\gamma$ is the probability measure with MG density process L_t such that

$$\begin{cases} dL_t = L_t \gamma_t \, dB_t \\ L_0 = 1 \end{cases} \tag{519}$$

Proof. By the definition of F, $\forall (\beta, \gamma) \in \mathscr{A}, f \le f^{\beta,\gamma}$ and there exists $(\hat{b}(t, y, z), \hat{c}(t, y, z))$ such that $f(t, y, z) = F(t, \hat{b}, \hat{c}) + y\hat{b} + z\hat{c}$. Since f is Lipschitz, $(\hat{b}, \hat{c})$ is bounded. As a result, there exists $(\hat{\beta}, \hat{\gamma})$ bounded progressive (measurable selection theorem, details not important here) such that $\forall t \in [0, T], f(t, Y_t, Z_t) = f^{\hat{\beta},\hat{\gamma}}(t, Y_t, Z_t)$. From the theorem above, we immediately conclude that

$$a.s. \; \forall t \in [0, T], Y_t = \inf_{(\beta,\gamma)\in\mathscr{A}} Y_t^{\beta,\gamma} \tag{520}$$

Now $(Y^{\beta,\gamma}, Z^{\beta,\gamma})$ is the solution to the linear BSDE

$$\begin{cases} dY_t = -[F(t, Y_t, Z_t) + Y_t\beta_t + Z_t\gamma_t] \, dt + Z_t \, dB_t \\ Y_T = \xi \end{cases} \tag{521}$$

so by previous theorems, we know that

$$\Gamma_t Y_t^{\beta,\gamma} = \mathbb{E}\left[\Gamma_T \xi + \int_t^T \Gamma_s F(s, Y_s, Z_s)\, ds \Big| \mathscr{F}_t\right] \tag{522}$$

for the adjoint process Γ_t as the solution to

$$\begin{cases} d\Gamma_t = \Gamma_t(\beta_t\, dt + \gamma_t\, dB_t) \\ \Gamma_0 = 1 \end{cases} \tag{523}$$

now set $\Gamma_t = e^{\int_0^t \beta_u\, du} L_t$ to find that $L_0 = 1$ and

$$d\Gamma_t = e^{\int_0^t \beta_u\, du} L_t \beta_t\, dt + e^{\int_0^t \beta_u\, du}\, dL_t \tag{524}$$

so L_t follows the SDE

$$dL_t = L_t \gamma_t\, dB_t \tag{525}$$

The last step comes from

$$\begin{aligned} Y_t^{\beta,\gamma} &= \mathbb{E}\left[\frac{\Gamma_T}{\Gamma_t}\xi + \int_t^T \frac{\Gamma_s}{\Gamma_t} F(s, Y_s, Z_s)\, ds \Big| \mathscr{F}_t\right] &(526)\\ &= \frac{1}{L_t}\mathbb{E}\left[\int_t^T e^{\int_t^s \beta_u\, du} F(s, \beta_s, \gamma_s) L_s\, ds + e^{\int_t^T \beta_u\, du}\xi L_T \Big| \mathscr{F}_t\right] &(527)\\ &= \mathbb{E}_{\mathbb{Q}^\gamma}\left[\int_t^T e^{\int_t^s \beta_u\, du} F(s, \beta_s, \gamma_s)\, ds + e^{\int_t^T \beta_u\, du}\xi \Big| \mathscr{F}_t\right] &(528) \end{aligned}$$

□

Remark. *Note that the **MG density process** here refers to the Radon-Nikodym derivative process restricted on a filtration as a MG. In other words, the probability measure $\mathbb{Q}^\gamma$ is given by*

$$\frac{d\mathbb{Q}^\gamma}{d\mathbb{P}}\Big|_{\mathscr{F}_t} = L_t \tag{529}$$

where the LHS is the Radon-Nikodym derivative of two probability measure restricted on sigma field $\mathscr{F}_t$.

*As a result, **the Bayes formula** can be written as: for $\mathbb{Q} << \mathbb{P}$ and Z_t as the MG density process of $\mathbb{Q}$ w.r.t. $\mathbb{P}$, for any stopping time $\sigma \leq \tau$, and $\xi \in \mathscr{F}_\tau$ such that $\mathbb{E}_{\mathbb{Q}}|\xi| < \infty$,*

$$\mathbb{E}_{\mathbb{Q}}(\xi|\mathscr{F}_\sigma) = \frac{\mathbb{E}_{\mathbb{P}}(Z_\tau \xi|\mathscr{F}_\sigma)}{Z_\sigma}\ \mathbb{Q} - a.s. \tag{530}$$

which can be proved easily by the definition of RN derivative that

$$\forall A \in \mathscr{F}_\sigma, \mathbb{E}_\mathbb{Q}[\mathbb{E}_\mathbb{Q}(\xi|\mathscr{F}_\sigma)Z_\sigma \mathbb{I}_A] = \mathbb{E}_\mathbb{Q}(\xi Z_\sigma \mathbb{I}_A) \tag{531}$$

$$\mathbb{E}_\mathbb{Q}[\mathbb{E}_\mathbb{P}(Z_\tau \xi|\mathscr{F}_\sigma)\mathbb{I}_A] = \mathbb{E}_\mathbb{P}[Z_\sigma \cdot \mathbb{E}_\mathbb{P}(Z_\tau \xi \mathbb{I}_A|\mathscr{F}_\sigma)] = \mathbb{E}_\mathbb{P}(Z_\sigma Z_\tau \xi \mathbb{I}_A) = \mathbb{E}_\mathbb{Q}(\xi Z_\sigma \mathbb{I}_A) \tag{532}$$

Take $\tau = \sigma = T$ and $Z = L$ to see that the last equation in the proof of the theorem above holds.

Remark. *To conclude, when the driver is concave, by taking the Frenchel conjugate F of driver f and varying the last two components of F as controls β, γ, one may find that $Y_t^{\beta,\gamma}$ is actually the problem value of a stochastic control problem under a different probability measure $\mathbb{Q}^\gamma$ with running cost*

$$\int_t^T e^{\int_t^s \beta_u \, du} F(s, \beta_s, \gamma_s) \, ds \tag{533}$$

in time interval $[t, T]$ and terminal cost

$$e^{\int_t^T \beta_u \, du} \xi \tag{534}$$

at time T (note that $\xi \in \mathscr{F}_T$). The concavity ensures that the double conjugate of f is still f itself and thus f has the structure as the inf of a family of linear drivers w.r.t. controls, leading to the solution to the BSDE Y_t as the inf of problem values, i.e. the value function.

Pontryagin Maximum Principle

The maximization formulation of stochastic control problem is slightly different from what we have done in the past. Here we consider the maximization problem instead of the minimization problem in finite time horizon. So the state process X_t still has dynamics

$$\begin{cases} dX_t = b(t, X_t, \alpha_t)\, dt + \sigma(t, X_t, \alpha_t)\, dB_t \\ X_0 = x \end{cases} \tag{535}$$

as a diffusion process with admissible control set $\mathscr{A}$ and the goal is to find optimal control α^* to achieve

$$\sup_{\alpha \in \mathscr{A}} \mathbb{E}\left[\int_0^T f(t, X_t, \alpha_t)\, dt + g(X_T)\right] \tag{536}$$

let's denote $J(\alpha) = \mathbb{E}\left[\int_0^T f(t, X_t, \alpha_t)\, dt + g(X_T)\right]$ as the problem value at time 0. All assumptions are the same to what we have made such that f, g are both nice enough, here we **require g to be C^1 and concave**.

We first define the **Hamiltonian** H as

$$H(t, x, \alpha, y, z) = b(t, x, \alpha) \cdot y + \sigma(t, x, \alpha) \cdot z + f(t, x, \alpha) \tag{537}$$

where $\cdot$ means the standard inner product and y, z are **dual variables** (one might have guessed that they are correspondent to (Y, Z) as the solution to BSDE). To be specific, for $y \in \mathbb{R}^d, z \in \mathbb{R}^{d\times m}$, $b(t, x, \alpha) \cdot y$ is the inner product between two vectors and $\sigma(t, x, \alpha)\cdot z$ is the inner product between two matrices defined as $\langle A, B\rangle = tr(A^T B)$. **Assume that $D_x H$ exists** and consider the **adjoint BSDE**

$$\begin{cases} dY_t = -D_x H(t, X_t, \alpha_t, Y_t, Z_t)\, dt + Z_t\, dB_t \\ Y_T = D_x g(X_T) \end{cases} \tag{538}$$

Theorem 23. ***(Pontryagin Maximum Principle)*** *Suppose that there exists $(\hat{Y}, \hat{Z})$ as the solution to the adjoint BSDE such that*

$$\forall t \in [0, T], H(t, \hat{X}_t, \hat{\alpha}_t, \hat{Y}_t, \hat{Z}_t) = \max_{\alpha \in \mathscr{A}} H(t, \hat{X}_t, \alpha, \hat{Y}_t, \hat{Z}_t) \tag{539}$$

where $\hat{X}_t$ is the solution to the SDE for given control $\hat{\alpha}$ and the adjoint BSDE is solved based on given $\hat{\alpha}, \hat{X}$. Now $H(t, x, \alpha, \hat{Y}_t, \hat{Z}_t)$ is almost surely concave in (x, α), then $\hat{\alpha}$ is an optimal control that maximizes $J(\alpha)$.

Proof. Only have to prove that $\forall \alpha \in \mathscr{A}, J(\hat{\alpha}) - J(\alpha) = \mathbb{E}\left[\int_0^T [f(t, \hat{X}_t, \hat{\alpha}_t) - f(t, X_t, \alpha_t)]\, dt + g(\hat{X}_T) - g(X_T)\right] \geq 0$. Tear into two parts to estimate the difference of the running cost and terminal cost respectively.

By the first-order condition of concavity,

$$\mathbb{E}\left[g(\hat{X}_T) - g(X_T)\right] \geq \mathbb{E}\left[D_x g(\hat{X}_T) \cdot (\hat{X}_T - X_T)\right] = \mathbb{E}\left[\hat{Y}_T \cdot (\hat{X}_T - X_T)\right] \tag{540}$$

by Ito formula, we can write the things inside the expectation as a sum of integrals

$$\mathbb{E}\left[\hat{Y}_T \cdot (\hat{X}_T - X_T)\right] = \mathbb{E}\left[\int_0^T (\hat{X}_t - X_t)\, d\hat{Y}_t + \int_0^T \hat{Y}_t\, d(\hat{X}_t - X_t) + \int_0^T d\left\langle Y, \hat{X} - X\right\rangle_t\right] \tag{541}$$

$$= \mathbb{E}\left[\int_0^T (\hat{X}_t - X_t)\, d\hat{Y}_t + \int_0^T \hat{Y}_t\, d(\hat{X}_t - X_t) + \int_0^T [\sigma(t, \hat{X}_t, \hat{\alpha}_t) - \sigma(t, X_t, \alpha_t)] \cdot \hat{Z}_t\, dt\right] \tag{542}$$

where $[\sigma(t, \hat{X}_t, \hat{\alpha}_t) - \sigma(t, X_t, \alpha_t)] \cdot \hat{Z}_t = tr\left([\sigma(t, \hat{X}_t, \hat{\alpha}_t) - \sigma(t, X_t, \alpha_t)]^T \hat{Z}_t\right)$ is the inner product. By **assuming that** (which can be verified after solving out everything)

$$\int_0^T (\hat{X}_t - X_t)\hat{Z}_t\, dB_t, \int_0^T \hat{Y}_t[\sigma(t, \hat{X}_t, \hat{\alpha}_t) - \sigma(t, X_t, \alpha_t)]\, dB_t \tag{543}$$

are both MGs in T, they won't contribute to the expectation and

$$\mathbb{E}\left[g(\hat{X}_T) - g(X_T)\right] \tag{544}$$

$$\geq \mathbb{E}\left[\int_0^T (\hat{X}_t - X_t)\, d\hat{Y}_t + \int_0^T \hat{Y}_t\, d(\hat{X}_t - X_t) + \int_0^T tr\left([\sigma(t, \hat{X}_t, \hat{\alpha}_t) - \sigma(t, X_t, \alpha_t)]^T \hat{Z}_t\right)\, dt\right] \tag{545}$$

$$= \mathbb{E}\Big[- \int_0^T (\hat{X}_t - X_t) \cdot D_x H(t, \hat{X}_t, \hat{\alpha}_t, \hat{Y}_t, \hat{Z}_t)\, dt + \int_0^T \hat{Y}_t[b(t, \hat{X}_t, \hat{\alpha}_t) - b(t, X_t, \alpha_t)]\, dt \tag{546}$$

$$+ \int_0^T tr\left([\sigma(t, \hat{X}_t, \hat{\alpha}_t) - \sigma(t, X_t, \alpha_t)]^T \hat{Z}_t\right)\, dt\Big] \tag{547}$$

For the running cost part, apply the definition of Hamiltonian to get

$$\mathbb{E}\left[\int_0^T [f(t, \hat{X}_t, \hat{\alpha}_t) - f(t, X_t, \alpha_t)]\, dt\right] \tag{548}$$

$$= \mathbb{E}\Big[\int_0^T [H(t, \hat{X}_t, \hat{\alpha}_t, \hat{Y}_t, \hat{Z}_t) - H(t, X_t, \alpha_t, \hat{Y}_t, \hat{Z}_t)]\, dt - \int_0^T \hat{Y}_t[b(t, \hat{X}_t, \hat{\alpha}_t) - b(t, X_t, \alpha_t)]\, dt \tag{549}$$

$$- \int_0^T tr\left([\sigma(t, \hat{X}_t, \hat{\alpha}_t) - \sigma(t, X_t, \alpha_t)]^T \hat{Z}_t\right)\, dt\Big] \tag{550}$$

Now take the sum of all those estimates to see

$$J(\hat{\alpha}) - J(\alpha) \geq \mathbb{E}\left[\int_0^T [H(t, \hat{X}_t, \hat{\alpha}_t, \hat{Y}_t, \hat{Z}_t) - H(t, X_t, \alpha_t, \hat{Y}_t, \hat{Z}_t) - (\hat{X}_t - X_t) \cdot D_x H(t, \hat{X}_t, \hat{\alpha}_t, \hat{Y}_t, \hat{Z}_t)]\, dt\right] \tag{551}$$

$$\geq 0 \tag{552}$$

by the conditions of the Pontryagin maximum principle. To see this, denote $H(t, \hat{X}_t, \hat{\alpha}_t, \hat{Y}_t, \hat{Z}_t)$ as $H(\hat{X}, \hat{\alpha})$ for convenience (only keep the X and α components). Assume that $D_\alpha H$ exists, then by joint concavity in (x, α),

$$H(\hat{X}, \hat{\alpha}) - H(X, \alpha) \geq D_x H(\hat{X}, \hat{\alpha}) \cdot (\hat{X} - X) + D_\alpha H(\hat{X}, \hat{\alpha}) \cdot (\hat{\alpha} - \alpha) \tag{553}$$

$$= D_x H(\hat{X}, \hat{\alpha}) \cdot (\hat{X} - X) \tag{554}$$

since $\hat{\alpha}$ is the control such that $H(\hat{X}, \alpha)$ is maximized, $D_\alpha H(\hat{X}, \hat{\alpha}) = 0$. For the general case, one can prove with subdifferential in the similar way. □

Remark. *There is a* ***connection between BSDE approach and PDE approach*** *that in the maximization problem the HJBE of value function $v(t, x)$ is formed as*

$$\partial_t v + \sup_{\alpha \in \mathscr{A}} \{G(t, x, \alpha, \partial_x v, \partial_{xx} v)\} \tag{555}$$

where $G(t, x, \alpha, \partial_x v, \partial_{xx} v) = L^\alpha v + f(t, x, \alpha)$.

When the value function is $C^{1,3}$ with nice enough regularity, and the existence of optimal control $\hat{\alpha} \in \mathscr{A}$ is ensured, with $\hat{X}$ as the state process generated by the dynamics for given control $\hat{\alpha}$, then

$$G(t, \hat{X}_t, \hat{\alpha}_t, \partial_x v(t, \hat{X}_t), \partial_{xx} v(t, \hat{X}_t)) = \max_{\alpha \in \mathscr{A}} G(t, \hat{X}_t, \alpha, \partial_x v(t, \hat{X}_t), \partial_{xx} v(t, \hat{X}_t)) \tag{556}$$

so the optimal strategy maximizes the sup in the HJBE when the state process is fixed as $\hat{X}$. Moreover,

$$(\hat{Y}_t, \hat{Z}_t) = (\partial_x v(t, \hat{X}_t), \partial_{xx} v(t, \hat{X}_t) \cdot \sigma(t, \hat{X}_t, \hat{\alpha}_t)) \tag{557}$$

is just the solution to the adjoint BSDE (similar to the form in the motivation of BSDE we have introduced above).

Remark. *In* ***minimization*** *problems, it's obvious that then we shall* ***minimize the Hamiltonian*** *instead of maximizing it and we will just require H* ***to be convex in*** *(x, α) instead of being concave, g* ***to be convex*** *instead of being concave.*

Stochastic Control Problem: BSDE Approach

By the Pontryagin maximum principle, we state the main steps of solving the stochastic control problem.

- **Get the Hamiltonian and solve for** $\hat{\alpha}(t, x, y, z) = \max_\alpha H(t, x, \alpha, y, z)$ for $\forall t, x, y, z$, this will give the control α as a function of t, x, y, z that always satisfies the condition in the Pontryagin maximum principle

- Solve the **coupled FBSDE** (Forward-Backward SDE)

$$\begin{cases} dX_t = b(t, X_t, \hat{\alpha}(t, X_t, Y_t, Z_t))\, dt + \sigma(t, X_t, \hat{\alpha}(t, X_t, Y_t, Z_t))\, dB_t \\ X_0 = x \\ dY_t = -D_x H(t, X_t, \hat{\alpha}(t, X_t, Y_t, Z_t), Y_t, Z_t)\, dt + Z_t\, dB_t \\ Y_T = D_x g(X_T) \end{cases} \tag{558}$$

  to get $\hat{\alpha}, \hat{X}, \hat{Y}, \hat{Z}$ as a set of solution. Note that we plug in the control variable as a function of t, X, Y, Z to apply Pontryagin maximum principle. The FSDE is the dynamics of the state process X while the BSDE is the dynamics of the adjoint (Y, Z). This FBSDE is called coupled since two SDEs cannot be solved separately.

- Verify the concavity of $H(t, x, \alpha, \hat{Y}_t, \hat{Z}_t)$ in (x, α) and the admissibility of $\hat{\alpha}$ as the optimal control. One might notice that the optimal control $\hat{\alpha}(t, \hat{X}_t, \hat{Y}_t, \hat{Z}_t)$ depends on the solution to the FBSDE. As a result, one has to first solve the FBSDE and then determine from measurability which kind of optimal control the solution corresponds to (open-loop, close-loop etc.).

One advantage of BSDE approach is that it makes it possible for us to find optimal controls other than Markovian controls. Note that the PDE approach depends on HJBE, which has its root in the dynamic programming principle with a natural Markovian structure. As a result, **the PDE approach can only be applied for the Markovian case**. On the other hand, as we will see in a later context, **the BSDE approach can be applied to find open-loop, closed-loop or Markovian optimal controls** which has much wider applications.

It might be quite obvious that the most difficult step lies in solving the coupled FBSDE. Let's use some examples to illustrate the BSDE approach.

Example: Exponential Utility Maximization with Option Payoff

Now there is a riskless asset with price $S^0 = 1$ that does not change with time, i.e. with no interest rate, and a risky asset in the market whose price S_t at time t follows the SDE

$$dS_t = S_t(b_t\,dt + \sigma_t\,dB_t) \tag{559}$$

where b, σ are bounded progressive processes with $\exists \varepsilon > 0, a.s.\ \forall t \in [0, T], \sigma_t \geq \varepsilon$. A person has wealth X_t at time t and the control variable is α_t that denotes the amount of wealth invested in risky asset at time t (similar setting to Merton problem). We know that the total wealth process follows the SDE

$$\begin{cases} dX_t = \alpha_t \frac{dS_t}{S_t} = \alpha_t(b_t\,dt + \sigma_t\,dB_t) \\ X_0 = x \end{cases} \tag{560}$$

with the admissible set of controls $\mathscr{A}$ as the collection of all progressive α such that $\int_0^T ||\alpha_t||^2\,dt < \infty\ a.s.$ and that the solution $X^{x,\alpha}$ for given α and initial value x is lower bounded.

Now a person has to **replicate the option payoff** ξ at time T for bounded $\xi \in \mathscr{F}_T$ (a European option but not necessarily a call or a put, the option can have any reasonable payoff function on the day of maturity) in order to hedge the risk contained in selling such an option. The utility function is now exponential and concave

$$U(x) = -e^{-\eta x}\ (\eta > 0) \tag{561}$$

so one's goal is to **find the optimal control to maximize the expected terminal utility under the condition that one fully hedges the risk of selling the option**. In other words, this person will have $X_T^{x,\alpha}$ wealth on the day of maturity following control $\alpha \in \mathscr{A}$ but since he is selling out this option, he will have to pay the option payoff ξ, so one wants to maximize

$$v(x) = \sup_{\alpha \in \mathscr{A}} \mathbb{E}U(X_T^{x,\alpha} - \xi) \tag{562}$$

A certain approach to deal with this problem lies in constructing a family of process $\{J_t^\alpha\}_{t\in[0,T]}, \alpha \in \mathscr{A}$ **such that**

- $\forall \alpha \in \mathscr{A}, J_T^\alpha = U(X_T^{x,\alpha} - \xi)$
- J_0^α **is constant and is independent of** $\alpha \in \mathscr{A}$
- $\forall \alpha \in \mathscr{A}, J^\alpha$ **is a super-MG and** $\exists \hat{\alpha} \in \mathscr{A}$ **such that** $J^{\hat{\alpha}}$ **is a MG**

The point of doing so is that now

$$\forall \alpha \in \mathscr{A}, \mathbb{E}U(X_T^{x,\alpha} - \xi) = \mathbb{E}J_T^\alpha \leq \mathbb{E}J_0^\alpha = \mathbb{E}J_0^{\hat{\alpha}} = \mathbb{E}J_T^{\hat{\alpha}} = \mathbb{E}U(X_T^{x,\hat{\alpha}} - \xi) \tag{563}$$

so $\hat{\alpha} \in \mathscr{A}$ is just the optimal control and $v(x) = \mathbb{E}J_0^{\hat{\alpha}}$ is the maximum expected utility.

Let's consider the family

$$J_t^\alpha = U(X_t^{x,\alpha} - Y_t) \tag{564}$$

here with (Y, Z) as the solution to the BSDE

$$\begin{cases} dY_t = -f(t, Z_t)\, dt + Z_t\, dB_t \\ Y_T = \xi \end{cases} \tag{565}$$

for the driver f to be specified later. Now the first two conditions of J_t^α are naturally satisfied with the value function $v(x) = U(x - Y_0)$ (note that now the filtration is taken as the Brownian filtration and the solution to the BSDE should be adapted, which means that $Y_0 \in \mathscr{F}_0$ so Y_0 is almost surely constant). To satisfy the third condition, we have to do some transformations to J_t^α to use its structure. Apply Ito formula to find

$$\begin{aligned}
\log(-J_t^\alpha) &= -\eta(X_t - Y_t) && (566)\\
&= -\eta(x - Y_0) - \eta \int_0^t dX_s + \eta \int_0^t dY_s && (567)\\
&= -\eta(x - Y_0) + \int_0^t (-\eta\alpha_s b_s - \eta f(s, Z_s))\, ds + \int_0^t (\eta Z_s - \eta\alpha_s\sigma_s)\, dB_s && (568)\\
J_t^\alpha &= -e^{-\eta(x-Y_0)} e^{\int_0^t (-\eta\alpha_s b_s - \eta f(s,Z_s))\, ds} e^{\int_0^t (\eta Z_s - \eta\alpha_s\sigma_s)\, dB_s} && (569)
\end{aligned}$$

now let's notice that the stochastic integral on the exponential can be written in the form as an exponential local MG that

$$J_t^\alpha = -e^{-\eta(x-Y_0)} e^{\int_0^t (\eta Z_s - \eta\alpha_s\sigma_s)\, dB_s - \frac{1}{2}\int_0^t (\eta Z_s - \eta\alpha_s\sigma_s)^2\, ds} e^{\int_0^t (-\eta\alpha_s b_s - \eta f(s,Z_s) + \frac{1}{2}(\eta Z_s - \eta\alpha_s\sigma_s)^2)\, ds} \tag{570}$$

now the product of first two terms (excluding the negative sign) on RHS form an non-negative exponential local MG M_t^α (so it must be a super-MG) and we just have to look at the last term

$$\begin{aligned}
C_t^\alpha &= -e^{\int_0^t (-\eta\alpha_s b_s - \eta f(s,Z_s) + \frac{1}{2}(\eta Z_s - \eta\alpha_s\sigma_s)^2)\, ds} && (571)\\
&= -e^{\eta \int_0^t \frac{\eta}{2}(Z_s - \alpha_s\sigma_s)^2 - \alpha_s b_s - f(s,Z_s)\, ds} && (572)\\
&= -e^{\eta \int_0^t \rho(s,\alpha_s,Z_s)\, ds} && (573)
\end{aligned}$$

where $\rho(t, a, z) = \frac{\eta}{2}(z - a\sigma_t)^2 - ab_t - f(t, z)$. **In order to make J_t^α a super-MG for $\forall \alpha \in \mathscr{A}$ and to make it a MG for some $\hat{\alpha} \in \mathscr{A}$, we have to make some restrictions on ρ order to specify the driver f in the BSDE**. One direct observation is that now J_t^α is already a product of C_t^α and a super-MG, so we want to see C_t^α

being decreasing in t **for** $\forall\alpha \in \mathscr{A}$ such that

$$\forall\alpha \in \mathscr{A}, \forall 0 < s < t, \mathbb{E}(J_t^\alpha|\mathscr{F}_s) = \mathbb{E}(M_t^\alpha C_t^\alpha|\mathscr{F}_s) \le C_s^\alpha \cdot \mathbb{E}(M_t^\alpha|\mathscr{F}_s) \le C_s^\alpha M_s^\alpha = J_s^\alpha \tag{574}$$

in order to ensure this, we only need to add the condition that

$$\forall\alpha \in \mathscr{A}, \forall t \in [0,T], \rho(t, \alpha_t, Z_t) \ge 0 \tag{575}$$

On the other hand, we want to make sure that $\exists\hat{\alpha} \in \mathscr{A}$ such that $J_t^{\hat{\alpha}} = M_t^{\hat{\alpha}} C_t^{\hat{\alpha}}$ is a MG. To let this be true, notice that we only need to ensure that $\exists\hat{\alpha} \in \mathscr{A}, C_t^\alpha = -1$, i.e. $C_t^{\hat{\alpha}}$ **is constantly** -1 **for some** $\hat{\alpha}$. Now it's still unclear why such condition helps us ensure that $J_t^{\hat{\alpha}}$ is a MG but we will see later (since it depends on the form of $\hat{\alpha}$). Let's add the condition that

$$\exists\hat{\alpha} \in \mathscr{A}, \forall t \in [0,T], \rho(t, \hat{\alpha}_t, Z_t) = 0 \tag{576}$$

Now the problem **turns into finding an appropriate driver** f **such that** $\rho(t, \alpha_t, Z_t) = \frac{\eta}{2}(Z_t - \alpha_t\sigma_t)^2 - \alpha_t b_t - f(t, Z_t)$ **is always non-negative and can reach 0 for some** $\hat{\alpha} \in \mathscr{A}$. Naturally, consider taking $\alpha_t = \frac{1}{\sigma_t}\left(Z_t + \frac{1}{\eta}\frac{b_t}{\sigma_t}\right)$ to minimize

$$\frac{\eta}{2}(Z_t - \alpha_t\sigma_t)^2 - \alpha_t b_t \tag{577}$$

and find that as a result we should **take the driver of BSDE as**

$$f(t, Z_t) = -Z_t\frac{b_t}{\sigma_t} - \frac{1}{2\eta}\frac{b_t^2}{\sigma_t^2} \tag{578}$$

The last step is to verify that $J_t^{\hat{\alpha}}$ is a MG for

$$\hat{\alpha}_t = \frac{1}{\sigma_t}\left(Z_t + \frac{1}{\eta}\frac{b_t}{\sigma_t}\right) \in \mathscr{A} \tag{579}$$

plug in to find

$$J_t^{\hat{\alpha}} = -M_t^{\hat{\alpha}} = -e^{-\eta(x - Y_0)} e^{\int_0^t (\eta Z_s - \eta\hat{\alpha}_s\sigma_s)\, dB_s - \frac{1}{2}\int_0^t (\eta Z_s - \eta\hat{\alpha}_s\sigma_s)^2\, ds} \tag{580}$$

ignoring the constant, it is the exponential local MG of $-\int_0^t \frac{b_s}{\sigma_s}\, dB_s$ where $\left\{\frac{b_s}{\sigma_s}\right\}_{s\in[0,T]} \in \mathbb{H}^{2,1}$ since b, σ are bounded process and σ is bounded away from 0. As a result, this exponential local MG is actually a true MG and we are done with all the conditions of J_t^α. The **optimal control** is just given by

$$\hat{\alpha}_t = \frac{1}{\sigma_t}\left(Z_t + \frac{1}{\eta}\frac{b_t}{\sigma_t}\right) \in \mathscr{A} \tag{581}$$

and the value function is

$$v(x) = -e^{-\eta(x-Y_0)} \tag{582}$$

where (Y, Z) is the solution to the BSDE

$$\begin{cases} dY_t = \left(Z_t \frac{b_t}{\sigma_t} + \frac{1}{2\eta}\frac{b_t^2}{\sigma_t^2}\right) dt + Z_t\, dB_t \\ Y_T = \xi \end{cases} \tag{583}$$

Remark. *Let's now talk about the intuition of this method. It's easily seen that this problem has something to do with* ***replicating a financial derivative in a complete financial market****. In a complete financial market, any contingent claim that pays bounded random payoff $\xi \in \mathscr{F}_T$ on the maturity date T can be* ***perfectly*** *replicated by a* ***self-financing*** *wealth process.*

On the other hand, this problem can also be understood in the risk-neutral world where risk-neutral measure $\mathbb{Q}$ is introduced such that

$$dS_t = S_t \sigma_t\, dB_t^{\mathbb{Q}} \tag{584}$$

so

$$B_t^{\mathbb{Q}} = B_t^{\mathbb{P}} + \int_0^t \frac{b_s}{\sigma_s}\, ds \tag{585}$$

where $\frac{d\mathbb{Q}}{d\mathbb{P}} = e^{-\int_0^T \frac{b_s}{\sigma_s}\, dB_s - \frac{1}{2}\int_0^T \frac{b_s^2}{\sigma_s^2}\, ds}$, now $x_\xi = \mathbb{E}_{\mathbb{Q}}\xi$ is just the ***no-arbitrage price of such option*** *and $\exists \pi \in \mathscr{A}, \xi = X_T^{x_\xi, \pi}$ as the* ***continuous-time Delta-hedging strategy****. The problem is equivalent to maximizing*

$$v(x) = \sup_{\alpha \in \mathscr{A}} \mathbb{E}_{\mathbb{P}} U(X_T^{x-x_\xi, \alpha-\pi}) \tag{586}$$

with 0 payoff received at time T (already hedged by π). They are actually two sides of the same coin. By calculations above, we know $\pi_t = \frac{Z_t}{\sigma_t}$ and the optimal control for this new problem to be $\hat{\alpha}_t = \frac{1}{\eta}\frac{b_t}{\sigma_t^2}$.

Remark. *One might be confused with the meaning of the Y, Z, J^α we mentioned above in the original problem. Here Y_t stands for the value of the European option evaluated at time t. That's why $Y_T = \xi$ is $\mathscr{F}_T$ measurable and is random (since it should be a function of S_T, the stock price on the maturity date) and Y_0 standing for the current price of this option is a constant.* ***BSDE has a natural correspondence with the evolution of option value*** *since option values are the easiest to evaluate on the maturity date but are difficult to figure out previous to maturity.*

J_t^α is just the utility at time t following control α in the position of selling out the option in this example, its meaning is actually the same as the "problem value" in the stochastic control problem and is closely connected with the value function. Here we adopt neither the dynamic programming approach nor the Pontryagin maximum principle but use another method to ***form the problem value J_t^α as a super-MG for all control and a MG for the***

optimal control. *Note that the martingality is always the hardest part to construct and shall always be verified base on the specific form of the optimal control (which means that one always has to "guess" some possible conditions such that the martingality holds, solve out the optimal control and go back to verify).*

One might be curious about the meaning of Z_t. *Now let's assume that* $Y_t = v(t, X_t)$ *since* Y_t *stands for the value of the option at time* t, *a natural structure of value function. By Ito formula,*

$$dY_t = \partial_t v(t, X_t)\, dt + \partial_x v(t, X_t)\, dX_t + \partial_{xx} v(t, X_t)\, d\langle X, X\rangle_t \tag{587}$$

assume that $dX_t = b(t, X_t)\, dt + \sigma(t, X_t)\, dB_t$ *and compare the* dB_t *term on both sides with the BSDE to get* $Z_t\, dB_t = \partial_x v(t, X_t)\sigma(t, X_t)\, dB_t$ *so* $Z_t = \partial_x v(t, X_t)\sigma(t, X_t)$, *which can be heuristically understood as the* ***Delta-Hedging strategy*** *since* $\partial_x v$ *is the sensitivity of option value w.r.t. stock price, the definition of option Delta.*

So far, we have built the understanding in BSDE that it describes the evolution of option value and Delta hedging strategy, it then makes sense that these two concepts shall be solved simultaneously (binomial option pricing).

When the market is incomplete, such method constructing J_t^α *still holds, but one may get a more complicated driver* f *(not linear in* Z_t *any longer).*

Example: Mean-Variance Criterion for Portfolio Selection

There's a riskless asset with price S_t^0 at time t with continuous-time interest rate r and a stock with price S_t at time t following the Black-Scholes model

$$dS_t = S_t(b\,dt + \sigma\,dB_t) \tag{588}$$

where $b > r, \sigma > 0$ are given constants. A person chooses to invest α_t amount of wealth in the stock at time t (the control), and he has total wealth X_t at time t, so we would be able to write down the dynamics of X_t

$$dX_t = [rX_t + \alpha_t(b - r)]\,dt + \sigma\alpha_t\,dB_t \tag{589}$$

with initial wealth $X_0 = x$ and the set of admissible controls is given by $\mathscr{A}$ containing all progressive processes α taking value in $\mathbb{R}$ such that $\mathbb{E}\int_0^T \alpha_t^2\,dt < \infty$. So far, the setting is the same as that for the Merton problem. However, our objective is not to maximize the utility at time T but to **minimize the variance of the terminal total wealth under the condition that the expectation of the terminal total wealth equals a given constant m (mean-variance criterion)**.

In other words, we want to find the optimal control that achieves the following inf for given m

$$V(m) = \inf_{\alpha\in\mathscr{A}} \{Var(X_T) : \mathbb{E}X_T = m\} \tag{590}$$

however, the stochastic control problem is now constrained to the condition $\mathbb{E}X_T = m$ so it's not easy to deal with. In order to remove such constraint, we consider the conjugate

$$\tilde{V}(\lambda) = \inf_{\alpha\in\mathscr{A}} \left\{\mathbb{E}(X_T - \lambda)^2\right\} \tag{591}$$

The reason we are using conjugate of this special form is to **transform the constrained control problem into an equivalent unconstrained problem**. To see the conjugate relationship, $\forall\varepsilon > 0, \exists\alpha^\varepsilon \in \mathscr{A}$ that generates the diffusion $X_t^{x,\alpha^\varepsilon}$ such that $\mathbb{E}X_T^{x,\alpha^\varepsilon} = m, Var(X_T^{x,\alpha^\varepsilon}) \le V(m)+\varepsilon$. So $\mathbb{E}(X_T^{x,\alpha^\varepsilon} - \lambda)^2 = Var(X_T^{x,\alpha^\varepsilon}) + \mathbb{E}^2(X_T^{x,\alpha^\varepsilon} - \lambda)$

$$\forall m, \forall\varepsilon > 0, \tilde{V}(\lambda) \le \mathbb{E}(X_T^{x,\alpha^\varepsilon} - \lambda)^2 \le V(m) + \varepsilon + (m - \lambda)^2 \tag{592}$$

similarly, $\forall\varepsilon > 0, \exists\alpha^\lambda \in \mathscr{A}$ that generates the diffusion X_t^{x,α^λ} such that $\mathbb{E}(X_T^{x,\alpha^\lambda} - \lambda)^2 \le \tilde{V}(\lambda) + \varepsilon$. So we have

$$\tilde{V}(\lambda) \ge \mathbb{E}(X_T^{x,\alpha^\lambda} - \lambda)^2 - \varepsilon = Var(X_T^{x,\alpha^\lambda}) + \mathbb{E}^2(X_T^{x,\alpha^\lambda} - \lambda) - \varepsilon \tag{593}$$

$$= V(\mathbb{E}X_T^{x,\alpha^\lambda}) + (\mathbb{E}X_T^{x,\alpha^\lambda} - \lambda)^2 - \varepsilon \tag{594}$$

as a result, we conclude that

$$\tilde{V}(\lambda) = \inf_{m\in\mathbb{R}} \left\{V(m) + (m - \lambda)^2\right\} \tag{595}$$

On the other hand, we also have

$$V(m) = \sup_{\lambda\in\mathbb{R}} \left\{\tilde{V}(\lambda) - (m-\lambda)^2\right\} \tag{596}$$

since $\tilde{V}(\lambda) = -2\sup_{m\in\mathbb{R}}\left\{m\lambda - \frac{V(m)+m^2}{2}\right\} + \lambda^2$ can be written as the form of Frenchel conjugate and $\frac{V(m)+m^2}{2}$ is strictly convex with Frenchel conjugate

$$\sup_{m\in\mathbb{R}}\left\{m\lambda - \frac{V(m)+m^2}{2}\right\} = \frac{\lambda^2 - \tilde{V}(\lambda)}{2} \tag{597}$$

so its double conjugate is itself and

$$\frac{V(m)+m^2}{2} = \sup_{m\in\mathbb{R}}\left\{m\lambda - \frac{\lambda^2 - \tilde{V}(\lambda)}{2}\right\} \tag{598}$$

which proves the equality above. So far, we have seen **the conjugate relationship between V and $\tilde{V}$**. This is important because this enables us to shift our gears from finding the optimal control achieving the inf in $V(m)$ to finding the optimal control achieving the inf in $\tilde{V}(\lambda)$. To see the connections between those two problems, if $\lambda = \lambda_m$ achieves sup in $V(m) = \sup_{\lambda\in\mathbb{R}}\left\{\tilde{V}(\lambda) - (m-\lambda)^2\right\}$ then we immediately know that m achieves the inf in $\tilde{V}(\lambda_m) = \inf_{n\in\mathbb{R}}\left\{V(n) + (n-\lambda_m)^2\right\}$ and m is also the unique real number that can achieve this sup by strict convexity. From the calculations we have done above, the inf in $\tilde{V}(\lambda_m)$ is achieved when $m = \mathbb{E}X_T^{x,\alpha^{\lambda_m}}$. As a result,

$$V(m) = Var(\mathbb{E}X_T^{x,\alpha^{\lambda_m}}) \tag{599}$$

so **the optimal control to the original problem is α^{λ_m}, the optimal control in the new problem setting with $\lambda = \lambda_m$ plugged in where λ_m is the λ that achieves the sup in $V(m) = \sup_{\lambda\in\mathbb{R}}\left\{\tilde{V}(\lambda) - (m-\lambda)^2\right\}$ such that the conjugacy holds**.

Now that we have transformed the problem into an easier one, our objective becomes finding optimal control that achieves the inf in

$$\tilde{V}(\lambda) = \inf_{\alpha\in\mathscr{A}}\left\{\mathbb{E}(X_T-\lambda)^2\right\} \tag{600}$$

for given $\lambda\in\mathbb{R}$. Recall the **Pontryagin maximum principle** that

$$H(t,x,a,y,z) = [rx + a(b-r)]y + \sigma a z \tag{601}$$

so $D_xH(t,x,a,y,z) = ry, g(x) = (x-\lambda)^2, D_xg(x) = 2(x-\lambda)$ and the adjoint BSDE is

$$\begin{cases} dY_t = -rY_t\,dt + Z_t\,dB_t \\ Y_T = 2(X_T-\lambda) \end{cases} \tag{602}$$

let $(\hat{Y},\hat{Z})$ be the solution pair to this BSDE, $\hat{\alpha}\in\mathscr{A}$ be the optimal control and $\hat{X}$ the diffusion generated by $\hat{\alpha}$, notice that

$$H(t,x,a,\hat{Y}_t,\hat{Z}_t)=[rx+a(b-r)]\hat{Y}_t+\sigma a\hat{Z}_t \tag{603}$$

is linear in (x,a) so it satisfies the concavity condition. Now we want to maximize the Hamiltonian $H(t,x,a,y,z)$ w.r.t. a to notice that this is possible if and only if

$$a.s.\ \forall t\in[0,T],(b-r)\hat{Y}_t+\sigma\hat{Z}_t=0 \tag{604}$$

by the linearity in a. The next step is to find the solution to the adjoint BSDE with the condition above also satisfied.

Here consider the **ansatz** (most often, one would like to consider the ansatz that $\hat{Y}_t$ is affine in $\hat{X}_t$ with the only randomness coming from $\hat{X}_t$)

$$\hat{Y}_t=\varphi(t)\hat{X}_t+\psi(t) \tag{605}$$

to get the ODEs for $\varphi,\psi\in C^1$ that

$$\begin{cases}\varphi'(t)\hat{X}_t+\varphi(t)[r\hat{X}_t+\hat{\alpha}_t(b-r)]+\psi'(t)+r[\varphi(t)\hat{X}_t+\psi(t)]=0\\ \varphi(t)\sigma\hat{\alpha}_t-\hat{Z}_t=0\\ \varphi(T)=2,\psi(T)=-2\lambda\end{cases} \tag{606}$$

the second ODE and the maximization of Hamiltonian give the optimal control

$$\hat{\alpha}_t=\frac{\hat{Z}_t}{\sigma\varphi(t)}=\frac{-\frac{b-r}{\sigma}\hat{Y}_t}{\sigma\varphi(t)}=\frac{(r-b)\hat{Y}_t}{\sigma^2\varphi(t)}=\frac{(r-b)[\varphi(t)\hat{X}_t+\psi(t)]}{\sigma^2\varphi(t)} \tag{607}$$

while the first ODE also gives the optimal control that

$$\hat{\alpha}_t=\frac{\varphi'(t)\hat{X}_t+\psi'(t)+r[\varphi(t)\hat{X}_t+\psi(t)]}{(r-b)\varphi(t)}+\frac{r}{r-b}\hat{X}_t=\frac{[\varphi'(t)+2r\varphi(t)]\hat{X}_t+\psi'(t)+r\psi(t)}{(r-b)\varphi(t)} \tag{608}$$

since those two optimal controls shall be the same for the original problem, by comparing the coefficients of $\hat{X}_t$ and the constant terms (since φ,ψ are deterministic), we get the simplified ODE systems for φ,ψ that

$$\begin{cases}\varphi'(t)=\left(\frac{(r-b)^2}{\sigma^2}-2r\right)\varphi(t)\\ \psi'(t)=\left(\frac{(r-b)^2}{\sigma^2}-r\right)\psi(t)\\ \varphi(T)=2,\psi(T)=-2\lambda\end{cases} \tag{609}$$

and they are very easy to solve

$$\begin{cases} \varphi(t) = 2e^{\left(\frac{(r-b)^2}{\sigma^2}-2r\right)(t-T)} \\ \psi(t) = -2\lambda e^{\left(\frac{(r-b)^2}{\sigma^2}-r\right)(t-T)} \end{cases} \tag{610}$$

Now we are done with the new stochastic control problem that achieves the inf in $\tilde{V}(\lambda)$. **The optimal Markovian control of the new problem** is given by

$$\forall\lambda \in \mathbb{R}, \hat{\alpha}_t^\lambda = \frac{(r-b)[\varphi(t)X_t + \psi_\lambda(t)]}{\sigma^2\varphi(t)} \tag{611}$$

note that **the dependence of λ only appears in ψ so we denote $\psi(t)$ as $\psi_\lambda(t)$**. To get the value function $\tilde{V}(\lambda)$, apply Ito formula for $\frac{1}{2}\varphi(t)X_t^2 + \psi_\lambda(t)X_t$ to see

$$X_T^2 - 2\lambda X_T = \frac{1}{2}\varphi(0)x^2 + \psi_\lambda(0)x + \int_0^T \frac{1}{2}\varphi'(t)X_t^2 + \psi_\lambda'(t)X_t\,dt \tag{612}$$

$$+\int_0^T \varphi(t)X_t + \psi_\lambda(t)\,dX_t + \frac{1}{2}\int_0^T \varphi(t)\,d\left\langle X, X\right\rangle_t \tag{613}$$

take an expectation on both sides and plug in the optimal control $\hat{\alpha}_t^\lambda$ to find that

$$\tilde{V}(\lambda) = e^{-\frac{(b-r)^2}{\sigma^2}T}(\lambda - e^{rT}x)^2 \tag{614}$$

The final step is to go back to the original problem. The value function $V(m)$ can be derived easily from the conjugate relationship. To get the optimal control $\hat{\alpha}_t^m$ for the original problem, recall that we only have to find λ_m that achieves the sup in $V(m) = \sup_{\lambda\in\mathbb{R}}\left\{\tilde{V}(\lambda) - (m-\lambda)^2\right\}$, some calculations tell us that

$$\lambda_m = \frac{m - e^{\left[r-\frac{(b-r)^2}{\sigma^2}\right]T}x}{1 - e^{-\frac{(b-r)^2}{\sigma^2}T}} \tag{615}$$

and **the optimal Markovian control for the original problem** is given by

$$\forall m \in \mathbb{R}, \hat{\alpha}_t^m = \frac{(r-b)[\varphi(t)X_t + \psi_{\lambda_m}(t)]}{\sigma^2\varphi(t)} \tag{616}$$

which ends the discussion.

Remark. *This example exhibits the way to apply Pontryagin maximum principle. Note that in this example we get no information on $\hat{\alpha}$ by maximizing the Hamiltonian (generally we would be able to represent $\hat{\alpha} = \hat{\alpha}(t, x, y, z)$ as a function and plug back to get FBSDE, which is more complicated). The **transformation to a new unconstrained stochastic control problem with conjugate value fucntion** is critical and greatly simplifies the calculations.*

Paper Summary: Mean Field Games and Systemic Risk

The following contents refers to the paper *Mean Field Games and Systemic Risk by Rene Carmona, Jean-Pierre Fouque and Li-Hsien Sun.*

In multi-agent games (finite number), we always want to find the **Nash Equilibrium (NE)**, which is a set of control that any one agent has no motivation to deviate from his own control given all other agents' controls. Let's consider a multi-agent control problem in the simple setting as an example.

For simplicity, we will consider the **linear-quadratic (LQ)** game with finitely many homogeneous players, for which there exists explicit solution of Nash Equilibrium.

Problem Setting

Now there are N banks (N large enough) in the economy as agents and X_t^i denotes the log-monetary reserves of the i-th bank at time t (state process) with the dynamics given as

$$dX_t^i = [a(\overline{X}_t - X_t^i) + \alpha_t^i]\,dt + \sigma(\sqrt{1-\rho^2}\,dB_t^i + \rho\,dB_t^0) \tag{617}$$

where $i = 1, 2, .., N$ denotes each bank and $B_t = (B_t^0, ..., B_t^N)$ is an $(N+1)$-dimensional BM. The reason we have the diffusion term $\sqrt{1-\rho^2}\,dB_t^i + \rho\,dB_t^0$ is that $\tilde{B}_t^i = \sqrt{1-\rho^2}B_t^i + \rho B_t^0$ gives N correlated BM. To see this, just compute the quadratic variation

$$\left\langle \tilde{B}^i, \tilde{B}^j \right\rangle_t = \left\langle \sqrt{1-\rho^2}B^i + \rho B^0, \sqrt{1-\rho^2}B^j + \rho B^0 \right\rangle_t = \begin{cases} t & i = j \\ \rho^2 t & i \neq j \end{cases} \tag{618}$$

to see that each $\tilde{B}^i$ is BM but they are not necessarily independent. This is the simple way we take to organize correlated BM using common noise B_t^0.

Here α_t^i is the control process for the i-th bank to determine. It can be explained as the borrowing/lending rate to a central bank. $a > 0$ is a given mean-reversion rate similar to that in the OU process and $\overline{X}_t = \frac{1}{N}\sum_{i=1}^N X_t^i$ always denotes the empirical mean.

Now each bank has its own problem value to **minimize**, the problem value of the i-th bank is

$$J^i(\alpha) = \mathbb{E}\left[\int_0^T f_i(X_t, \alpha_t^i)\,dt + g(X_T)\right] \tag{619}$$

note that since we are in the multi-agent setting, the problem value depends on $\alpha = (\alpha^1, ..., \alpha^N)$ not only its own control but also the control of other banks (since they jointly affect the evolution of state process X_t). The running cost depends on f_i which is a function of the state process $X_t = (X_t^1, .., X_t^N)$ and its own control α_t^i. The terminal cost only depends on the terminal state of the i-th bank. Now f_i, g_i has simple forms given by

$$\begin{cases} f_i(x,\alpha^i) = \frac{1}{2}(\alpha^i)^2 - q\alpha^i(\overline{x} - x^i) + \frac{\varepsilon}{2}(\bar{x} - x^i)^2 \\ g_i(x) = \frac{c}{2}(\overline{x} - x^i)^2 \end{cases} \tag{620}$$

quadratic functions, $q > 0$ controls the incentive of borrowing and lending from central bank and $c > 0$ is a parameter penalizing the deviation from the average state. Now we assume that $q^2 \le \varepsilon$ so f_i is convex in (x, α^i). The convexity will make sense in a later context.

For the convenience of notations, let's rewrite the whole dynamics of X_t in the vector form that

$$d\begin{bmatrix} X_t^1 \\ \cdots \\ X_t^N \end{bmatrix} = b(t, X_t, \alpha_t)\,dt + \sigma(t, X_t)\begin{bmatrix} dB_t^0 \\ dB_t^1 \\ \cdots \\ dB_t^N \end{bmatrix} \tag{621}$$

where

$$b(t, X_t, \alpha_t) = \begin{bmatrix} a(\overline{X}_t - X_t^1) + \alpha_t^1 \\ \cdots \\ a(\overline{X}_t - X_t^N) + \alpha_t^N \end{bmatrix} \in \mathbb{R}^N \tag{622}$$

$$\sigma(t, X_t) = \sigma\begin{bmatrix} \rho & \sqrt{1-\rho^2} & 0 & \cdots & 0 \\ \rho & 0 & \sqrt{1-\rho^2} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \rho & 0 & 0 & \cdots & \sqrt{1-\rho^2} \end{bmatrix} \in \mathbb{R}^{N\times(N+1)} \tag{623}$$

Markovian Case: PDE Approach

The first step of PDE approach is to build up the value function. Since now there are N banks, each bank has its own value function

$$V^i(t,x) = \inf_{\alpha\in\mathscr{A}} \mathbb{E}_{t,x}\left[\int_t^T f_i(X_t,\alpha_t^i)\,dt + g_i(X_T^i)\right]\ (t\in\mathbb{R}, x\in\mathbb{R}^N) \tag{624}$$

where $\mathbb{E}_{t,x}$ means that the expectation is under the initial value condition that $X_t = x$ and $\mathscr{A}$ is the admissible set with integrability conditions for all **Markovian controls (here it means that $\mathcal{I}_t = \sigma(t, X_t)$ so when making the decision each bank has complete information on all banks' states)**. Note that since the problem value for the i-th banks depends on the state of all banks (f_i is a function of $X_t^1, ..., X_t^N$), here we must include all the states of the system in the value function. The HJBE for finite time horizon tells us that

$$\begin{cases} \partial_t V^i + \inf_{\alpha^i\in\mathscr{A}}\left\{L^\alpha V + f_i\right\} = 0 \\ V^i(T,x) = g_i(x) \end{cases} \tag{625}$$

with the action of the infinitesimal generator

$$L^\alpha V = \sum_{j=1}^N [a(\overline{x}-x^j)+\alpha^j]\partial_{x^j}V^i \tag{626}$$

$$+\frac{\sigma^2}{2}tr\left(\begin{bmatrix}\rho & \sqrt{1-\rho^2} & 0 & \cdots & 0\\ \rho & 0 & \sqrt{1-\rho^2} & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ \rho & 0 & 0 & \cdots & \sqrt{1-\rho^2}\end{bmatrix}\begin{bmatrix}\rho & \sqrt{1-\rho^2} & 0 & \cdots & 0\\ \rho & 0 & \sqrt{1-\rho^2} & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ \rho & 0 & 0 & \cdots & \sqrt{1-\rho^2}\end{bmatrix}^T H\right) \tag{627}$$

$$= \sum_{j=1}^N [a(\overline{x}-x^j)+\alpha^j]\partial_{x^j}V^i + \frac{\sigma^2}{2}\sum_{j,k=1}^N[\rho^2+\delta_{j,k}(1-\rho^2)]\partial_{x^j,x^k}V^i \tag{628}$$

now for player i, he can only determine which α^i to take and all other players' controls are actually functions of t, x denoted

$$\alpha^j = \phi^j(t,x)\ (j\neq i) \tag{629}$$

at this point we can write down the **HJBE** as

$$\begin{cases} \partial_t V^i + \inf_{\alpha^i\in\mathscr{A}}\left\{\sum_{j=1}^N [a(\overline{x}-x^j)+\alpha^j]\partial_{x^j}V^i + \frac{\sigma^2}{2}\sum_{j,k=1}^N[\rho^2+\delta_{j,k}(1-\rho^2)]\partial_{x^j,x^k}V^i + f_i\right\} = 0 \\ V^i(T,x) = g_i(x) \end{cases} \tag{630}$$

First get rid of the inf by considering the optimization problem

$$\min_{\alpha^i} Q(\alpha) = \sum_{j=1}^N [a(\overline{x} - x^j) + \alpha^j]\partial_{x^j} V^i + \frac{\sigma^2}{2}\sum_{j,k=1}^N [\rho^2 + \delta_{j,k}(1-\rho^2)]\partial_{x^j,x^k} V^i + f_i \tag{631}$$

notice that actually only the terms relevant to α^i matters

$$\begin{aligned} \frac{\partial Q}{\partial \alpha^i} &= \partial_{x^i} V^i + \frac{\partial f_i}{\partial \alpha^i} &(632)\\ &= \partial_{x^i} V^i + \alpha^i - q(\overline{x} - x^i) = 0 &(633)\end{aligned}$$

the optimal Markovian control is given by

$$\hat{\alpha}^i_t = q(\overline{X}_t - X^i_t) - \partial_{x^i} V^i(t, X_t) \tag{634}$$

where $\partial_{x^i} V^i$ is unknown and needs to be solved. Since this is the optimal control for the i-th bank and in this game symmetricity holds, **every bank shall have this optimal Markovian control to form Markovian NE**.

Now plug back into the HJBE to see that it becomes

$$\begin{aligned} &\partial_t V^i + \sum_{j=1}^N [(a+q)(\overline{x} - x^j) - \partial_{x^j} V^j]\partial_{x^j} V^i + \frac{\sigma^2}{2}\sum_{j,k=1}^N [\rho^2 + \delta_{j,k}(1-\rho^2)]\partial_{x^j,x^k} V^i &(635)\\ &+ \frac{1}{2}(\varepsilon - q^2)(\overline{x} - x^i)^2 + \frac{1}{2}(\partial_{x^i} V^i)^2 = 0 &(636)\end{aligned}$$

use the **ansatz** that

$$V^i(t,x) = \frac{\eta_t}{2}(\overline{x} - x^i)^2 + \mu_t \tag{637}$$

for deterministic functions η_t, μ_t to get the terminal conditions $\eta_T = c, \mu_T = 0$ and

$$\begin{cases} \partial_{x^j} V^i = \eta_t(\overline{x} - x^i)\left(\frac{1}{N} - \delta_{i,j}\right) \\ \partial_{x^j,x^k} V^i = \eta_t\left(\frac{1}{N} - \delta_{i,k}\right)\left(\frac{1}{N} - \delta_{i,j}\right) \end{cases} \tag{638}$$

the HJBE now becomes

$$\begin{aligned} &\frac{\eta_t'}{2}(\overline{x} - x^i)^2 + \mu_t' + \sum_{j=1}^N \left[a + q - \eta_t\left(\frac{1}{N} - 1\right)\right]\eta_t(\overline{x} - x^i)(\overline{x} - x^j)\left(\frac{1}{N} - \delta_{i,j}\right) &(639)\\ &+ \frac{\sigma^2}{2}\sum_{j,k=1}^N [\rho^2 + \delta_{j,k}(1-\rho^2)]\eta_t\left(\frac{1}{N} - \delta_{i,k}\right)\left(\frac{1}{N} - \delta_{i,j}\right) &(640)\\ &+ \frac{1}{2}(\varepsilon - q^2)(\overline{x} - x^i)^2 + \frac{1}{2}\left[\eta_t(\overline{x} - x^j)\left(\frac{1}{N} - 1\right)\right]^2 = 0 &(641)\end{aligned}$$

where the only randomness comes from the term $(\overline{x}-x^i)^2$, so its coefficient has to be zero, and the constant terms shall also add up to zero. We get two ODEs where η_t', μ_t' are the derivatives w.r.t. time t. Note that here $\sum_{j=1}^N(\overline{x}-x^j)\left(\frac{1}{N}-\delta_{i,j}\right) = -(\overline{x}-x^i)$, a simple transformation

$$\frac{\eta_t'}{2} - \left[a+q-\eta_t\left(\frac{1}{N}-1\right)\right]\eta_t + \frac{1}{2}(\varepsilon-q^2) + \frac{1}{2}\left[\eta_t\left(\frac{1}{N}-1\right)\right]^2 = 0 \tag{642}$$

$$\mu_t' + \frac{\sigma^2}{2}\sum_{j,k=1}^N[\rho^2+\delta_{j,k}(1-\rho^2)]\eta_t\left(\frac{1}{N}-\delta_{i,k}\right)\left(\frac{1}{N}-\delta_{i,j}\right) = 0 \tag{643}$$

Some simplifications tell us that

$$\begin{cases} \eta_t' = \left(1-\frac{1}{N^2}\right)\eta_t^2 + 2(a+q)\eta_t - (\varepsilon-q^2) \\ \mu_t' = -\frac{\sigma^2(1-\rho^2)}{2}\left(1-\frac{1}{N}\right)\eta_t \\ \eta_T = c, \mu_T = 0 \end{cases} \tag{644}$$

to get a **Ricatti equation** (first-order ODE with the derivative η_t' equal to a quadratic in η_t) with constant coefficients for η_t and another equation for μ_t which is easily solve as

$$\mu_t = \frac{\sigma^2(1-\rho^2)}{2}\left(1-\frac{1}{N}\right)\int_t^T \eta_s\, ds \tag{645}$$

if η_t is known. Now the problem only lies on solving out η_t, note that luckily this Ricatti equation is separable in variables so we only have to use the integral that

$$\int \frac{1}{Ax^2+Bx+C}\,dx = \frac{1}{A}\int\frac{1}{y^2-\frac{B^2-4AC}{4A^2}}\,dy \;\left(y = x+\frac{B}{2A}\right) \tag{646}$$

$$= \frac{1}{A}\int\frac{1}{\left(y-\frac{\sqrt{\Delta}}{2A}\right)\left(y+\frac{\sqrt{\Delta}}{2A}\right)}\,dy \tag{647}$$

$$= \frac{1}{\sqrt{\Delta}}\left(\int\frac{1}{y-\frac{\sqrt{\Delta}}{2A}}\,dy - \int\frac{1}{y+\frac{\sqrt{\Delta}}{2A}}\,dy\right) \tag{648}$$

$$= \frac{1}{\sqrt{\Delta}}\left(\log\left(y-\frac{\sqrt{\Delta}}{2A}\right) - \log\left(y+\frac{\sqrt{\Delta}}{2A}\right)\right) \tag{649}$$

$$= \frac{1}{\sqrt{\Delta}}\left(\log\left(x+\frac{B}{2A}-\frac{\sqrt{\Delta}}{2A}\right) - \log\left(x+\frac{B}{2A}+\frac{\sqrt{\Delta}}{2A}\right)\right) \tag{650}$$

assuming that $A>0, \Delta = B^2-4AC>0$ ($q^2\le\varepsilon$ ensures that $B^2-4AC>0$ for our problem). So plug in all the coefficients and the terminal condition to get

$$\eta_t = \frac{-(\varepsilon - q^2)(e^{(\delta^+-\delta^-)(T-t)} - 1) - c(\delta^+ e^{(\delta^+-\delta^-)(T-t)} - \delta^-)}{(\delta^- e^{(\delta^+-\delta^-)(T-t)} - \delta^+) - c(1-\frac{1}{N^2})(e^{(\delta^+-\delta^-)(T-t)} - 1)} \tag{651}$$

where the $\delta^\pm, R$ are given by

$$\begin{cases} R = (a+q)^2 + (1-\frac{1}{N^2})(\varepsilon - q^2) > 0 \\ \delta^\pm = -(a+q) \pm \sqrt{R} \end{cases} \tag{652}$$

note that $\Delta = 4R, \delta^\pm = \frac{B\pm\sqrt{\Delta}}{2}$ and we keep the same notation as that in the paper.

For the last **verification** step which is omitted in the paper, notice that we have to make sure that the solution to HJBE is the value function and the admissibility of the optimal control. Recall that $V^i(t,x) = \frac{\eta_t}{2}(\overline{x} - x^i)^2 + \mu_t$ so the V^i solved meets the quadratic growth condition in x, $V^i \in C^{1,2}$. As a result, by checking the admissibility of optimal control, the verification theorem for finite-time horizon tells us that V^i is just the value function of the i-th bank. Now

$$\hat{\alpha}_t^i = q(\overline{X}_t - X_t^i) - \partial_{x^i} V^i(t, X_t) \tag{653}$$

$$= q(\overline{X}_t - X_t^i) - \left(\frac{1}{N} - 1\right)\eta_t(\overline{X}_t - X_t^i) \tag{654}$$

is measurable w.r.t. $\sigma(X_t^1, ..., X_t^N)$ so it's Markovian and the Nash equilibrium is solved out, the i-th bank will take the optimal Markovian strategy $\hat{\alpha}_t^i$ knowing the state of all other banks.

Markovian Case: BSDE Approach

Let's present another way to solve for the optimal Markovian control under Nash equilibrium, which is by the BSDE approach using the Pontryagin maximum principle. The first fact to notice is that in the state dynamics, the control only appears in the drift coefficient but not the diffusion coefficient and the diffusion coefficient also does not contain the state, so we can use **the reduced Hamiltonian** of the i-th bank

$$H^i(t,x,\alpha,y^{i,1},...,y^{i,N}) = \sum_{k=1}^{N}[a(\overline{x}-x^k)+\alpha^k]y^{i,k}+f_i(x,\alpha^i) \tag{655}$$

this is because we only care about minimizing the Hamiltonian w.r.t. the control and the adjoint BSDE only has something to do with $\partial_x H$. We have to claim here that since there are N banks in this game, the Hamiltonian has something to do with all the states of those banks and for each bank there exists dual variables $y^{i,1},...,y^{i,N}$ denoting the value of the process $(Y_t^{i,1},...,Y_t^{i,N})$ for the i-th bank (recall that the drift coefficient in the SDE for the state process is an N-dimensional vector, that's why for each bank there are N dual variables $y^{i,k}$ and the adjoint BSDE is a BSDE in N dimension).

However, since we are in the Markovian setting with complete information, we know that each bank can access the states of all banks $X_t^1,...,X_t^N$ at time t. As a result, when considering reaching Nash equilibrium, the i-th bank has to fix the control of all other banks, but **the control of all other banks shall be viewed as functions of t,x since we know that all of them are actually choosing feedback strategies based on time and the current state they are facing**. In other words, we are still acting as if we are in the perspective of the i-th bank and denote our strategy as α_t^i and denote other banks' strategy as $\alpha_t^k = \alpha_t^k(t,x)$ for $k \neq i$.

As a result, we shall write the Hamiltonian in the form that

$$H^i(t,x,\alpha,y^{i,1},...,y^{i,N}) = \sum_{k\neq i}[a(\overline{x}-x^k)+\alpha^k(t,x)]y^{i,k}+[a(\overline{x}-x^i)+\alpha^i]y^{i,i}+f_i(x,\alpha^i) \tag{656}$$

and minimize H^i w.r.t. variable α_t^i to get

$$\hat{\alpha}_t^i = q(\overline{x}-x^i)-y^{i,i} \tag{657}$$

since there's symmetricity across all the banks, the i-th bank has enough reason to believe that all other banks shall take the same optimal control

$$\hat{\alpha}^k(t,x) = q(\overline{x}-x^k)-y^{k,k}(t,x)\ (k\neq i) \tag{658}$$

the Hamiltonian's derivative w.r.t. variable x^j is

$$\partial_{x^j}H^i = a\left(\frac{1}{N}-\delta_{i,j}\right)y^{i,i}+\sum_{k\neq i}y^{i,k}\left[a\left(\frac{1}{N}-\delta_{j,k}\right)+\partial_{x^j}\alpha^k(t,x)\right]-q\alpha^i\left(\frac{1}{N}-\delta_{i,j}\right)+\varepsilon(\overline{x}-x^i)\left(\frac{1}{N}-\delta_{i,j}\right) \tag{659}$$

Now consider the **ansatz** that $Y_t^{i,j} = \eta_t \left(\frac{1}{N} - \delta_{i,j}\right)(\overline{X}_t - X_t^i)$ (the ansatz is often taken as an **affine function** for some transformations of X), plug in the formula for optimal control $\hat{\alpha}^k$ to find that

$$\begin{cases} \hat{\alpha}^k(t,x) = [q + \eta_t \left(1 - \frac{1}{N}\right)](\overline{x} - x^k) \\ \partial_{x^j} \hat{\alpha}^k(t,x) = \left[q + \eta_t \left(1 - \frac{1}{N}\right)\right]\left(\frac{1}{N} - \delta_{j,k}\right) \end{cases} \tag{660}$$

so the coefficient of dt in the adjoint BSDE is just $-\partial_{x^j} H^i(t, X_t, \hat{\alpha}_t, Y_t^{i,1}, ..., Y_t^{i,N})$. Let's plug in the optimal control for all agents and plug in the ansatz to get

$$\partial_{x^j} H^i(t, X_t, \hat{\alpha}_t, Y_t^{i,1}, ..., Y_t^{i,N}) = -\left(\frac{1}{N} - \delta_{i,j}\right)(\overline{X}_t - X_t^i)\left[-\frac{1}{N}\left(\frac{1}{N} - 1\right)\eta_t^2 + (a+q)\eta_t - (\varepsilon - q^2)\right] \tag{661}$$

with

$$\partial_{x^j} g_i(x) = c(\overline{x} - x^i)\left(\frac{1}{N} - \delta_{i,j}\right) \tag{662}$$

so **the adjoint BSDE** is given by

$$\begin{cases} dY_t^{i,j} = \left(\frac{1}{N} - \delta_{i,j}\right)(\overline{X}_t - X_t^i)\left[-\frac{1}{N}\left(\frac{1}{N} - 1\right)\eta_t^2 + (a+q)\eta_t - (\varepsilon - q^2)\right] dt + \sum_{k=0}^N Z_t^{i,j,k}\, dB_t^k \\ Y_T^{i,j} = c(\overline{X}_T - X_T^i)\left(\frac{1}{N} - \delta_{i,j}\right) \end{cases} \tag{663}$$

where $Z_t^{i,j,k}$ is the dual process for the i-th bank correspondent to dual process $Y_t^{i,j}$. Note that since we are having $N+1$ BM $B_t^0, ..., B_t^N$, k can take values $0, 1, ..., N$.

Now that we get the adjoint BSDE set up, it's still necessary that we plug in the optimal control and the ansatz in the state dynamics to get **a coupled FSDE**

$$dX_t^i = [(a+q)(\overline{X}_t - X_t^i) - Y_t^{i,i}]\, dt + \sigma[\sqrt{1-\rho^2}\, dB_t^i + \rho\, dB_t^0] \tag{664}$$

$$= (\overline{X}_t - X_t^i)\left[a + q - \left(\frac{1}{N} - 1\right)\eta_t\right] dt + \sigma[\sqrt{1-\rho^2}\, dB_t^i + \rho\, dB_t^0] \tag{665}$$

to conclude, the **coupled FBSDE** for this control problem is (already with ansatz plugged in)

$$\begin{cases} dX_t^i = (\overline{X}_t - X_t^i)\left[a + q - \left(\frac{1}{N} - 1\right)\eta_t\right] dt + \sigma[\sqrt{1-\rho^2}\, dB_t^i + \rho\, dB_t^0] \\ dY_t^{i,j} = \left(\frac{1}{N} - \delta_{i,j}\right)(\overline{X}_t - X_t^i)\left[-\frac{1}{N}\left(\frac{1}{N} - 1\right)\eta_t^2 + (a+q)\eta_t - (\varepsilon - q^2)\right] dt + \sum_{k=0}^N Z_t^{i,j,k}\, dB_t^k \\ Y_T^{i,j} = c(\overline{X}_T - X_T^i)\left(\frac{1}{N} - \delta_{i,j}\right) \end{cases} \tag{666}$$

note that for the i-th bank, we have $N+1$ unknown processes $X_t^i, Y_t^{i,1}, ..., Y_t^{i,N}$ and $N+1$ SDEs.

One canonical **strategy** to solve this FBSDE after plugging in the ansatz is **to take derivative for the ansatz, to plug in the FBSDE and to compare the coefficients with the BSDE**. Let's first take derivative w.r.t. t

for the ansatz

$$dY_t^{i,j} = \left(\frac{1}{N} - \delta_{i,j}\right) \eta_t \, d(\overline{X}_t - X_t^i) + \left(\frac{1}{N} - \delta_{i,j}\right) (\overline{X}_t - X_t^i) \eta_t' \, dt \tag{667}$$

it's natural that we want to replace $d(\overline{X_t} - X_t^i)$ with the known state dynamics. In order to do this, let's first take the sum of dX_t^i w.r.t. i to get

$$d\overline{X}_t = \sigma \left[\sqrt{1-\rho^2} \frac{1}{N} \sum_{k=1}^{n} dB_t^k + \rho \, dB_t^0\right] \tag{668}$$

compute the difference $d\overline{X}_t - dX_t^i$ to see

$$d(\overline{X}_t - X_t^i) = -(\overline{X}_t - X_t^i) \left[a + q - \left(\frac{1}{N} - 1\right) \eta_t\right] dt + \sigma \sqrt{1-\rho^2} \left[\frac{1}{N} \sum_{k=1}^{n} dB_t^k - dB_t^i\right] \tag{669}$$

therefore, we know that

$$dY_t^{i,j} = \left(\frac{1}{N} - \delta_{i,j}\right) (\overline{X}_t - X_t^i) \left[\eta_t' - \left(1 - \frac{1}{N}\right) \eta_t^2 - (a+q)\eta_t\right] dt + \sigma \sqrt{1-\rho^2} \eta_t \left(\frac{1}{N} - \delta_{i,j}\right) \sum_{k=1}^{N} \left(\frac{1}{N} - \delta_{i,k}\right) dB_t^k \tag{670}$$

by comparing the coefficients with the BSDE, we immediately know that

$$\begin{cases} Z^{i,j,0} = 0 \\ Z^{i,j,k} = \sigma\sqrt{1-\rho^2}\eta_t \left(\frac{1}{N} - \delta_{i,j}\right)\left(\frac{1}{N} - \delta_{i,k}\right) \ (k = 1, 2, ..., N) \\ \eta_t' = \left(1 - \frac{1}{N^2}\right)\eta_t^2 + 2(a+q)\eta_t - (\varepsilon - q^2) \\ \eta_T = c \end{cases} \tag{671}$$

one might verify that the process $Z^{i,j,k}$ is adapted and $Z^{i,j,k} \in \mathbb{H}^2$ since it's actually deterministic, and η_t satisfies the same ODE with the same initial value condition as that in the HJB approach, so the result is actually the same.

For the **verification** step left over, note that by Pontryagin maximum principle we just have to check that g_i is convex in x and that the Hamiltonian H^i with Y, Z as the solution to the BSDE plugged in, is convex in (x, α). The convexity of g is very easy to see from its Hessian matrix. On the other hand, fixing Y, Z in H^i, one will find that the Hamiltonian is convex in (x, α) since the first part $\sum_{k=1}^N [a(\overline{x} - x^k) + \alpha^k] y^{i,k}$ is linear in (x, α) and we have mentioned that $f_i(x, \alpha^i)$ is convex when $q^2 \leq \varepsilon$. So the sufficiency holds and now we see **the reason we are requiring that** $q^2 \leq \varepsilon$.

Open-loop Case: BSDE Approach

Now let's derive the open-loop optimal control for the same problem using BSDE approach. Note that for the open-loop case, the PDE approach deriving HJBE fails and Pontryagin maximum principle will be the most powerful tool. **The definition of open-loop control** is that we are requiring that $\alpha_t = \alpha_t(t, X_0, B_{[0,t]})$. In other words, the control taken at time t may depend on time, initial state and the value of all random noises from time 0 to t. The difference from Markovian case lies in the fact that **each agent now cannot observe all the agents' state evolution so feedback strategy is not allowed**.

Of course, the reduced Hamiltonian for the i-th bank is still

$$H^i(t, x, \alpha, y^{i,1}, ..., y^{i,N}) = \sum_{k=1}^{N} [a(\overline{x} - x^k) + \alpha^k] y^{i,k} + f_i(x, \alpha^i) \tag{672}$$

and here we just **take** $\alpha^1, ..., \alpha^N$ **as independent variables** since open loop NE provides no feedback effect. Minimize H^i w.r.t. α^i to get open loop NE

$$\hat{\alpha}_t^i = q(\overline{x} - x^i) - y^{i,i} \tag{673}$$

and naturally it holds for all agents.

Note that now $\hat{\alpha}^k$ is not a function in (t, x) any longer, so the computation of $\partial_{x^j} H^i$ gets simpler **(this is actually the difference between open-loop and Markovian cases)**

$$\partial_{x^j} H^i = \sum_{k=1}^{N} a \left(\frac{1}{N} - \delta_{j,k} \right) y^{i,k} - q\alpha^i \left(\frac{1}{N} - \delta_{i,j} \right) + \varepsilon(\overline{x} - x^i) \left(\frac{1}{N} - \delta_{i,j} \right) \tag{674}$$

let's plug in the optimal controls to get

$$\partial_{x^j} H^i = \frac{a}{N} \sum_{k=1}^{N} (y^{i,k} - y^{i,j}) + \left(\frac{1}{N} - \delta_{i,j} \right) (\varepsilon - q^2)(\overline{x} - x^i) + q \left(\frac{1}{N} - \delta_{i,j} \right) y^{i,i} \tag{675}$$

Now take the same **ansatz** as that in the Markovian case that $Y_t^{i,j} = \phi_t \left(\frac{1}{N} - \delta_{i,j} \right) (\overline{X}_t - X_t^i)$ to find that

$$\partial_{x^j} H^i(t, X_t, \hat{\alpha}_t, Y_t^{i,1}, ..., Y_t^{i,N}) = -\left(\frac{1}{N} - \delta_{i,j} \right) (\overline{X}_t - X_t^i) \left[a\phi_t + q \left(1 - \frac{1}{N} \right) \phi_t - (\varepsilon - q^2) \right] \tag{676}$$

so **the adjoint BSDE** is given by

$$\begin{cases} dY_t^{i,j} = \left(\frac{1}{N} - \delta_{i,j} \right) (\overline{X}_t - X_t^i) \left[a\phi_t + q \left(1 - \frac{1}{N} \right) \phi_t - (\varepsilon - q^2) \right] \, dt + \sum_{k=0}^{N} Z_t^{i,j,k} \, dB_t^k \\ Y_T^{i,j} = c(\overline{X}_T - X_T^i) \left(\frac{1}{N} - \delta_{i,j} \right) \end{cases} \tag{677}$$

Now we get **coupled FBSDE** (the derivation of FSDE is exactly the same as that in the Markovian case, so we neglect all details here)

$$\begin{cases} dX_t^i = (\overline{X}_t - X_t^i)\left[a + q - \left(\frac{1}{N} - 1\right)\phi_t\right]\, dt + \sigma[\sqrt{1-\rho^2}\, dB_t^i + \rho\, dB_t^0] \\ dY_t^{i,j} = \left(\frac{1}{N} - \delta_{i,j}\right)(\overline{X}_t - X_t^i)\left[a\phi_t + q\left(1 - \frac{1}{N}\right)\phi_t - (\varepsilon - q^2)\right]\, dt + \sum_{k=0}^N Z_t^{i,j,k}\, dB_t^k \\ Y_T^{i,j} = c(\overline{X}_T - X_T^i)\left(\frac{1}{N} - \delta_{i,j}\right) \end{cases} \tag{678}$$

which is a little bit different from that for the Markovian case. However, the trick applied to solve it is exactly the same, by comparing the coefficients of the derivative of the ansatz with the BSDE, we get (refer to the Markovian case for detailed calculations, exactly the same)

$$\begin{cases} Z^{i,j,0} = 0 \\ Z^{i,j,k} = \sigma\sqrt{1-\rho^2}\phi_t\left(\frac{1}{N} - \delta_{i,j}\right)\left(\frac{1}{N} - \delta_{i,k}\right) \ (k = 1, 2, ..., N) \\ \phi_t' = \left(1 - \frac{1}{N}\right)\phi_t^2 + 2\left[a + q\left(1 - \frac{1}{2N}\right)\right]\phi_t - (\varepsilon - q^2) \\ \phi_T = c \end{cases} \tag{679}$$

one might verify that the process $Z^{i,j,k}$ is adapted and $Z^{i,j,k} \in \mathbb{H}^2$ since it's actually deterministic, and ϕ_t satisfies the ODE with initial value condition. We won't solve explicitly this Ricatti equation once more here, but it's an easy task since the ODE is separable in variables and we have already derived a general formula above for the integration (in the PDE approach part).

The **verification** part is still all about convexity and are verified in the Markovian case, so sufficiency holds and we are done. One extra thing to notice is that

$$\hat{\alpha}_t^i = \left[q + \left(1 - \frac{1}{N}\right)\phi_t\right](\overline{X}_t - X_t^i) \tag{680}$$

the optimal control looks like a Markovian control. However, since we have the dynamics for $\overline{X}_t - X_t^i$ which is

$$d(\overline{X}_t - X_t^i) = -(\overline{X}_t - X_t^i)\left[a + q - \left(\frac{1}{N} - 1\right)\eta_t\right]\, dt + \sigma\sqrt{1-\rho^2}\left[\frac{1}{N}\sum_{k=1}^n dB_t^k - dB_t^i\right] \tag{681}$$

we can solve out $\overline{X}_t - X_t^i$ and represent it as a function of X_0 and $B_{[0,t]}$ (to see the explicit expression, since the diffusion coefficient is constant, we can first ignore the diffusion term and then change the constant into a process to solve this SDE).

The following materials refer to *Spectral Graph Theory by Fan Chung*.

Spectral Graph Theory

Before entering the stochastic control problems on graphs, let's first introduce important elements in spectral graph theory that will help us simplify the calculations. The spectral graph theory focuses on the Laplacian matrix of the graph and the eigenvalues of such Laplacian matrix since it will tell us about the structures of the graphs.

Graph Laplacian

The most important concept is the **Laplacian of a graph** $G = (V, E)$ where V is the vertex set and E is the edge set. Let's assume that G is **undirected and simple**, denote d_v as the degree of vertex $v \in V$, then for vertices $\forall u, v \in V$, define the matrix L as

$$L(u, v) = \begin{cases} d_v & u = v \\ -1 & u \sim v \\ 0 & else \end{cases} \tag{682}$$

where $u \sim v$ means that there exists an edge between u and v. Define the matrix T as

$$T = diag(\{d_v\}_{v \in V}) \tag{683}$$

to be a diagonal matrix with the degree of all vertices on the diagonal. The Laplacian $\mathcal{L}$ is defined as

$$\mathcal{L} = T^{-\frac{1}{2}} L T^{-\frac{1}{2}} \tag{684}$$

so one may find that $\forall u, v \in V$

$$\mathcal{L}(u, v) = \begin{cases} 1 & u = v, d_v \neq 0 \\ -\frac{1}{\sqrt{d_u d_v}} & u \sim v \\ 0 & else \end{cases} \tag{685}$$

and we set the convention $T^{-1}(v, v) = 0$ if $d_v = 0$ for isolated vertices. One might notice the relationship between L, T and adjacent matrix A that $L = T - A$ so

$$\mathcal{L} = T^{-\frac{1}{2}}(T - A)T^{-\frac{1}{2}} = I - T^{-\frac{1}{2}} A T^{-\frac{1}{2}} = SS^T \tag{686}$$

for $S \in \mathbb{R}^{|V| \times |E|}$ constructed in a way that for each column corresponding to an edge $(u, v) \in E$, S has entry $\frac{1}{\sqrt{d_u}}$ on the row corresponding to u and has entry $-\frac{1}{\sqrt{d_v}}$ on the row corresponding to v. This is telling us that $\mathcal{L}$ is **symmetric and semi-positive-definite**, denoted $\mathcal{L} \in \mathbb{S}_+^n$ so it has real and non-negative eigenvalues.

An important interpretation of $\mathcal{L}$ lies in the fact that it's actually **an operator acting on graph functions**. Consider $g : V \to \mathbb{R}$ as a graph function, then

$$\mathcal{L}g(u) = \frac{1}{\sqrt{d_u}} \sum_{v:u\sim v} \left(\frac{g(u)}{\sqrt{d_u}} - \frac{g(v)}{\sqrt{d_v}}\right) \tag{687}$$

since $\mathcal{L} \in \mathbb{R}^{|V|\times|V|}$ can be seen as a matrix and $g \in \mathbb{R}^{|V|}$ can be seen as a vector so the action $\mathcal{L}g(u)$ can actually be understood as matrix multiplication for fixed $v \in V$.

Variational Characterization of Eigenvalues

For the eigenvalues of symmetric matrices, one characterization is to use the **Rayleigh quotient**, i.e. consider

$$\frac{\langle g, \mathcal{L}g\rangle}{\langle g, g\rangle} \tag{688}$$

where $g : V \to \mathbb{R}$ can be formed as a vector with dimension $|V|$ and $\langle\cdot,\cdot\rangle$ denoting the standard inner product. As a result, simple calculations tell us

$$\begin{aligned}
\frac{\langle g, \mathcal{L}g\rangle}{\langle g, g\rangle} = \frac{\langle f, Lf\rangle}{\langle f, Tf\rangle} &= \frac{\sum_v d_v f^2(v) - \sum_v f(v)\sum_{u:u\sim v} f(u)}{\sum_v d_v f^2(v)} && (689)\\
&= \frac{\sum_v \sum_{u:u\sim v} f^2(v) - \sum_{u\sim v} 2f(u)f(v)}{\sum_v d_v f^2(v)} && (690)\\
&= \frac{\sum_{u\sim v}[f^2(u) + f^2(v)] - \sum_{u\sim v} 2f(u)f(v)}{\sum_v d_v f^2(v)} && (691)\\
&= \frac{\sum_{u\sim v}[f(u) - f(v)]^2}{\sum_v d_v f^2(v)} && (692)
\end{aligned}$$

by assuming that $g = T^{\frac{1}{2}} f$ and noticing that $\sum_v \sum_{u:u\sim v}$ will visit each adjacent pair (u, v) for 2 times while $\sum_{u\sim v}$ will visit each adjacent pair (u, v) for 1 time.

The Rayleigh quotient expression tells us immediately that **the graph Laplacian $\mathcal{L}$ has eigenvalue 0**. Denote all eigenvalues as $0 \le \lambda_0 \le \lambda_1 \le ... \le \lambda_{n-1}$, forming the **spectrum** of Laplacian, then

$$\lambda_0 = \inf_{g\in\mathbb{R}^{|V|}} \frac{\langle g, \mathcal{L}g\rangle}{\langle g, g\rangle} = \inf_{f\in\mathbb{R}^{|V|}} \frac{\sum_{u\sim v}[f(u) - f(v)]^2}{\sum_v d_v f^2(v)} = 0 \tag{693}$$

the inf is achieved by f such that $\forall u \sim v, f(u) = f(v)$. One example may be

$$f = \vec{1}, f(v) = 1 \tag{694}$$

where $\vec{1}$ denotes a column vector with all entries to be 1 so $g = T^{\frac{1}{2}}\vec{1}$ **is the eigenfunction of $\mathcal{L}$ for eigenvalue 0**.

Remark. *Here we have to use* ***Courant's minmax characterization of eigenvalues*** *for a real symmetric matrix* *that the j-th smallest eigenvalue is*

$$\lambda_{j-1} = \min_{dimS=j} \max_{g\in S, g\neq 0} \frac{\langle g, \mathcal{L}g\rangle}{\langle g, g\rangle} \tag{695}$$

$$= \inf_{f\perp TP_{j-1}} \frac{\sum_{u\sim v}[f(u)-f(v)]^2}{\sum_v d_v f^2(v)} \tag{696}$$

where S is a linear subspace of $\mathbb{R}^{|V|}$ and P_{j-1} is the eigenspace spanned by eigenfunctions correspondent to eigenvalues $\lambda_0, ..., \lambda_{j-1}$.

By such variational characterization, the second smallest eigenvalue has the characterization that

$$\lambda_1 = \min_{dimS=2} \max_{T^{\frac{1}{2}} f\in S, f\neq 0} \frac{\sum_{u\sim v}[f(u)-f(v)]^2}{\sum_v d_v f^2(v)} \tag{697}$$

$$= \inf_{f\perp T\vec{1}} \frac{\sum_{u\sim v}[f(u)-f(v)]^2}{\sum_v d_v f^2(v)} \tag{698}$$

and the largest eigenvalue is

$$\lambda_{n-1} = \sup_f \frac{\sum_{u\sim v}[f(u)-f(v)]^2}{\sum_v d_v f^2(v)} \tag{699}$$

Examples

Now let's compute the eigenvalues of some graphs that have good symmetricity as examples. The first example is the **complete graph** K_n, where

$$\mathcal{L} = \begin{bmatrix} 1 & \frac{1}{1-n} & \frac{1}{1-n} & ... & \frac{1}{1-n} \\ \frac{1}{1-n} & 1 & \frac{1}{1-n} & ... & \frac{1}{1-n} \\ ... & & & & \\ \frac{1}{1-n} & \frac{1}{1-n} & \frac{1}{1-n} & ... & 1 \end{bmatrix} \tag{700}$$

so from $\det(\lambda I - \mathcal{L}) = 0$, we get that (all columns have same sum, so we can transform the matrix into a diagonal one to calculate det)

$$\det \begin{bmatrix} \lambda - 1 & \frac{1}{n-1} & ... & \frac{1}{n-1} \\ \frac{1}{n-1} & \lambda - 1 & ... & \frac{1}{n-1} \\ ... & & & \\ \frac{1}{n-1} & \frac{1}{n-1} & ... & \lambda - 1 \end{bmatrix} = \lambda \left(\lambda - 1 - \frac{1}{n-1}\right)^{n-1} = 0 \tag{701}$$

so besides the trivial eigenvalue $\lambda_0 = 0$, we know that $\frac{n}{n-1}$ is also an eigenvalue with multiplicity $n-1$.

For the **complete bipartite graph** $K_{m,n}$ with m vertices on a side, n vertices on the other side, where

$$\mathcal{L} = \begin{bmatrix} I_m & -\frac{1}{\sqrt{mn}}E_{m\times n} \\ -\frac{1}{\sqrt{mn}}E_{n\times m} & I_n \end{bmatrix} \tag{702}$$

where E is the matrix with all 1 entries. It's obvious that 1 is an eigenvalue and $rank(I-\mathcal{L}) = 2$ so it has multiplicity $m+n-2$, besides the eigenvalue 0, there's one more multiplicity left for another eigenvalue. By observation,

$$\mathcal{L}\begin{bmatrix} -\sqrt{n} \\ ... \\ -\sqrt{n} \\ \sqrt{m} \\ ... \\ \sqrt{m} \end{bmatrix} = 2\begin{bmatrix} -\sqrt{n} \\ ... \\ -\sqrt{n} \\ \sqrt{m} \\ ... \\ \sqrt{m} \end{bmatrix} \tag{703}$$

since $-\sqrt{n} - \frac{n\sqrt{m}}{\sqrt{mn}} = -2\sqrt{n}, \frac{m\sqrt{n}}{\sqrt{mn}} + \sqrt{m} = 2\sqrt{m}$. As a result, the eigenvalues are 0 with multiplicity 1, 1 with multiplicity $m+n-2$ and 2 with multiplicity 1.

For the **star graph** S_n, it's defined as the graph with one vertex with degree $n-1$ in the middle and other $n-1$ vertices with degree 1 around. Let's assume that v_1 is the vertex in the middle, so

$$\mathcal{L} = \begin{bmatrix} 1 & -\frac{1}{\sqrt{n-1}} & -\frac{1}{\sqrt{n-1}} & ... & -\frac{1}{\sqrt{n-1}} \\ -\frac{1}{\sqrt{n-1}} & 1 & 0 & ... & 0 \\ ... & & & & \\ -\frac{1}{\sqrt{n-1}} & 0 & 0 & ... & 1 \end{bmatrix} \tag{704}$$

so by Laplacian expansion of determinant (assume the determinant is of size $k \times k$, to distinguish from n in the recurrence relationship)

$$D_k = \det(\lambda I - \mathcal{L}) = (\lambda-1)D_{k-1} + (-1)^{n+1}\frac{1}{\sqrt{n-1}}(-1)^n\frac{1}{\sqrt{n-1}}(\lambda-1)^{k-2} \tag{705}$$

$$= (\lambda-1)D_{k-1} - \frac{(\lambda-1)^{k-2}}{n-1} \tag{706}$$

$$\frac{D_k}{(\lambda-1)^{k-2}} = (\lambda-1)D_1 - \frac{k-1}{n-1} = (\lambda-1)^2 - \frac{k-1}{n-1} \tag{707}$$

now plug in $k=n$ to see that

$$D_n = (\lambda-1)^{n-2}\lambda(\lambda-2) \tag{708}$$

so star graph has eigenvalues 0 with multiplicity 1, 1 with multiplicity $n-2$, 2 with multiplicity 1.

For the **path graph** P_n where all vertices are in a single path $v_1 \to v_2 \to ... \to v_n$, v_1, v_n has degree 1 while all other vertices have degree 2, so the Laplacian is

$$\mathcal{L} = \begin{bmatrix} 1 & -\frac{1}{\sqrt{2}} & 0 & 0 & 0 & ... & 0 & 0 & 0 \\ -\frac{1}{\sqrt{2}} & 1 & -\frac{1}{2} & 0 & 0 & ... & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 1 & -\frac{1}{2} & 0 & ... & 0 & 0 & 0 \\ ... \\ 0 & 0 & 0 & 0 & 0 & ... & -\frac{1}{2} & 1 & -\frac{1}{\sqrt{2}} \\ 0 & 0 & 0 & 0 & 0 & ... & 0 & -\frac{1}{\sqrt{2}} & 1 \end{bmatrix} \tag{709}$$

which is a tri-diagonal determinant. For general tri-diagonal determinant with $a_1, ..., a_n$ on the diagonal, $b_1, ..., b_{n-1}$ on super-diagonal and $c_1, ..., c_{n-1}$ on sub-diagonal, the determinant has recurrence relationship

$$D_k = a_k D_{k-1} - c_{k-1} b_{k-1} D_{k-2} \tag{710}$$

with $D_0 = 1, D_{-1} = 0$. So for $D_n = \det(\lambda I - \mathcal{L})$, we see that $\forall 1 \le i \le n, a_i = \lambda - 1, b_1 = c_1 = b_{n-1} = c_{n-1} = \frac{1}{\sqrt{2}}, \forall 2 \le i \le n-2, b_i = c_i = \frac{1}{2}$

$$\begin{cases} D_n = (\lambda - 1)D_{n-1} - \frac{1}{2}D_{n-2} \\ D_{n-1} = (\lambda - 1)D_{n-2} - \frac{1}{4}D_{n-3} \\ ... \\ D_3 = (\lambda - 1)D_2 - \frac{1}{4}D_1 \\ D_2 = (\lambda - 1)D_1 - \frac{1}{2}D_0 \end{cases} \tag{711}$$

and the eigenvalues are given by $1 - \cos\frac{k\pi}{n-1}$ $(k = 0, 1, ..., n-1)$. (too much calculations, not verified yet)

For the **cycle graph** C_n, each vertex has degree 2 and the Laplacian is

$$\mathcal{L} = \begin{bmatrix} 1 & -\frac{1}{2} & 0 & 0 & 0 & ... & 0 & 0 & -\frac{1}{2} \\ -\frac{1}{2} & 1 & -\frac{1}{2} & 0 & 0 & ... & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 1 & -\frac{1}{2} & 0 & ... & 0 & 0 & 0 \\ ... \\ 0 & 0 & 0 & 0 & 0 & ... & -\frac{1}{2} & 1 & -\frac{1}{2} \\ -\frac{1}{2} & 0 & 0 & 0 & 0 & ... & 0 & -\frac{1}{2} & 1 \end{bmatrix} \tag{712}$$

a circulant matrix. As a result, the eigenvalues are directly given by

$$\lambda_j = 1 - \frac{1}{2}\omega^j - \frac{1}{2}\omega^{(n-1)j} \ (j = 0, 1, ..., n-1) \tag{713}$$

where $\omega = e^{\frac{2\pi i}{n}}$ is the n-th complex root. After simplification, it's not hard to see that

$$\lambda_j = 1 - \cos \frac{2\pi j}{n} \ (j = 0, 1, ..., n-1) \tag{714}$$

Property of the Graph Spectrum

Let's only prove some simple properties for the graph spectrum since that would suffice for our purpose of looking into stochastic control problems on graphs.

Theorem 24. ***(General Results)*** *It's always true that*

$$\sum_i \lambda_i \leq n \tag{715}$$

the equality holds iff G has no isolated vertices. For $n \geq 2$,

$$\lambda_1 \leq \frac{n}{n-1} \tag{716}$$

the equality holds iff $G = K_n$. For G without isolated vertices,

$$\lambda_{n-1} \geq \frac{n}{n-1} \tag{717}$$

If $G \neq K_n$, $\lambda_1 \leq 1$.

Proof. The sum of eigenvalues is just $tr(\mathcal{L})$ so it's obvious that the trace is no larger than n and reaches n iff all vertices have at least 1 degree.

Now for $n \geq 2$, $\lambda_0 = 0$ so $n \geq \lambda_1 + ... + \lambda_{n-1} \geq (n-1)\lambda_1$ gives the trivial bound. When the equality is true, there's no isolated vertices and $\lambda_1 = \lambda_2 = ... = \lambda_{n-1} = \frac{n}{n-1}$, same as the spectrum of K_n (note that there's one-to-one correspondence between graph spectrum and the graph). If G has no isolated vertices, $n = \lambda_1 + ... + \lambda_{n-1} \leq (n-1)\lambda_{n-1}$, so $\lambda_{n-1} \geq \frac{n}{n-1}$.

If $G \neq K_n$ then there exists $a, b \in V, a \not\sim b$, consider graph function

$$f_1(v) = \begin{cases} d_b & v = a \\ -d_a & v = b \\ 0 & else \end{cases} \tag{718}$$

so $f \perp T\vec{1}$ and

$$\lambda_1 = \inf_f \frac{\sum_{u \sim v}[f(u) - f(v)]^2}{\sum_v d_v f^2(v)} \tag{719}$$

$$\leq \frac{(d_a - 1)d_b^2 + (d_b - 1)d_a^2}{d_b d_a^2 + d_a d_b^2} \leq 1 \tag{720}$$

$\square$

For the connectedness of a graph, it is also reflected in the spectrum in the following results.

Theorem 25. ***(Connectedness)*** *If G is connected, then $\lambda_1 > 0$. If $\lambda_i = 0, \lambda_{i+1} \neq 0$, then G has exactly $i+1$ connected components.*

$$\forall i \leq n-1, \lambda_i \leq 2 \tag{721}$$

with $\lambda_{n-1} = 2$ iff a connected component of G is bipartite and nontrivial.

Generally, the spectrum of G is the union of the spectrum of its connected components.

Proof. Different connected components in G correspond to disjoint blocks on the diagonal of $\mathcal{L}$ (after reordering the vertices according to which connect component they are in), that's why it's easy to see that the union of the spectrum of connected components is the spectrum of G.

So if the graph is not connected, there's at least 2 connected components and eigenvalue 0 has at least multiplicity 2. Conversely, if the graph is connected, then if $\lambda_1 = \inf_{f \perp T\vec{1}} \frac{\sum_{u\sim v}[f(u)-f(v)]^2}{\sum_v d_v f^2(v)} = 0$ then the f that can achieve this inf is constant on all vertices, which is the same as the eigenfunction for λ_0, a contradiction, so $\lambda_1 > 0$. As a result, we immediately see that a graph has $i+1$ connected components iff the spectrum of the graph is the union of the spectrum of $i+1$ connected graphs iff $\lambda_0 = ... = \lambda_i = 0, \lambda_{i+1} > 0$.

To set up an upper bound for the eigenvalue, notice that

$$[f(u) - f(v)]^2 \leq 2f^2(u) + 2f^2(v) \tag{722}$$

so $\lambda_{n-1} = \sup_f \frac{\sum_{u\sim v}[f(u)-f(v)]^2}{\sum_v d_v f^2(v)} \leq 2\sup_f \frac{\sum_{u\sim v}[f^2(u)+f^2(v)]}{\sum_v d_v f^2(v)} = 2$. The equality holds iff the function $f \neq 0$ that achieves the sup is such that $\forall u \sim v, f(u) = -f(v)$. If the equality holds, then such f exists, so consider a nontrivial connected component of G (since there exists non-zero eigenvalue, such component must exist) on which f is not constantly 0 (since $f \neq 0$). On such connected component there's only two possible values $a, -a \neq 0$ taken by f, collect all vertices of this connected component taking value a under f as V_1 and collect all vertices of this connected component taking value $-a$ under f as V_2, then edges only exists between V_1 and V_2 but not within V_1 or V_2, so it's a bipartite connected component. Conversely, if there exists a bipartite connected component with vertex sets partitioned into V_1, V_2, then $\forall a > 0, \forall v \in V_1, f(v) = a, \forall v \in V_2, f(v) = -a$, such f suffices to achieve the sup. $\square$

Theorem 26. ***(Bipartite Graph)*** *G is bipartite iff G has i connected components and $\forall 1 \leq j \leq i, \lambda_{n-j} = 2$ iff $\forall \lambda_i, 2 - \lambda_i$ is also en eigenvalue of G.*

Proof. WLOG, only prove for the case where G is connected. From the last theorem, it's bipartite iff the largest eigenvalue is 2, proved.

Now there exists eigenfunction f corresponding to λ_i so $\lambda_i = \frac{\sum_{u\sim v}[f(u)-f(v)]^2}{\sum_v d_v f^2(v)}$, if G is bipartite, the vertex set can be partitioned into V_1, V_2, so consider

$$g(x) = \begin{cases} f(x) & x \in V_1 \\ -f(x) & x \in V_2 \end{cases} \tag{723}$$

with $\frac{\sum_{u\sim v}[g(u)-g(v)]^2}{\sum_v d_v g^2(v)} = \frac{\sum_{u\sim v}[f(u)+f(v)]^2}{\sum_v d_v f^2(v)} = 2\frac{\sum_{u\sim v}[f^2(u)+f^2(v)]}{\sum_v d_v f^2(v)} - \frac{\sum_{u\sim v}[f(u)-f(v)]^2}{\sum_v d_v f^2(v)} = 2-\lambda_i$. g is just the eigenfunction corresponding to eigenvalue $2-\lambda_i$ is G is bipartite. Conversely, since $\lambda_0 = 0$, $\lambda_{n-1} = 2$ so G has to be bipartite and we have proved the theorem. □

Remark. *Intuitively,* ***the multiplicity of eigenvalue 0 is equal to the number of connected components of*** G ***and the multiplicity of eigenvalue 2 is equal to the number of bipartite connected components of*** G *since for bipartite connected component, its eigenvalues always appear in a pair that adds up to 2.*

We only introduce those elements in spectral graph theory so far and we will introduce more important properties if necessary.

The contents below refers to the paper *The Graph Neural Network Model by Franco Scarselli, Marco Gori et. al..*

Graph Neural Network (GNN)

Overview

I want to put the contents that has something to do with GNN here since it provides a new perspective to look at stochastic control problems on graphs. In the graph learning field, the practice in the past is to use the vector embedding of a graph, i.e. turning the graphical information as a vector and only deal with that vector. This idea is often adopted in the machine learning field like the word embedding in NLP turning words into its vector embeddings. This approach is called network representation learning (NRL) focusing on how to represent a network well in different problem settings to maintain as much information as possible.

However, GNN takes another approach to accept the whole graph as part of the input of a neural network, avoiding the representation step and tries to keep all information in the graphical structure. GNN can deal with graphs with directed or undirected edges (even the mixture), positional or non-positional graphs so it's generally applicable to making predictions with graphical structures or approximating functions on graphs.

Basic Setting

Graph learning problems always fall into one of the two categories: **graph-focused** or **node-focused**. The former one aims at approximating a function $\tau : \mathscr{G} \to \mathbb{R}^m$ where $\mathscr{G}$ is the set of all graphs we are interested in, i.e. the value of τ has nothing to do with each node in the graph, e.g. predicting the probability a chemical causes a disease based on its molecule structure. On the other hand, if we denote $G = (N, E) \in \mathscr{G}$ as a graph with node set N and edge set E, we want to approximate $\tau : \mathscr{G} \times N \to \mathbb{R}^m$ in node-focused tasks. Such τ may change it value if different node is taken in the graph, e.g. predicting the impact factor of a person on social network, then the celebrities are more likely to have a higher impact factor than normal people.

Again, GNN deals with both kinds of tasks under the same framework, and the main idea is **graph connectivity**, i.e. the neighbors of each node in the graph (of course it's possible to consider a more general definition of k-neighbor where path of length less or equal to k exists). Let

$$ne[n] = \{m \in N : m \sim n\} \tag{724}$$

denote the neighborhood of node n and

$$co[n] = \{e \in E : \exists m \in N, e = (m, n)\} \tag{725}$$

denote the set of edges having n as one endpoint. Here we adopt the notations as if the graph is undirected.

All information w.r.t. nodes and edges is contained in the **labels**. For node n, $l_n \in \mathbb{R}^{l_N}$ denotes its label and for edge $e = (n_1, n_2)$, $l_{(n_1,n_2)} \in \mathbb{R}^{l_E}$ denotes its label. Labels are formed as real-valued vectors carrying information of the nodes and edges. The label is actually one of the attributes that makes GNN work in the general context. To be specific, we can add one digit in the label of the edge to specify if the edge is directed, which enables us to consider graphs with a mixture of directed and undirected edges. We can also add to the label the logical position of the nodes so that it can deal with positional graphs. In short, the label enables us to encode a lot of information and greatly increases the generality of the model.