SELECTED TOPICS IN CONTROL THEORY
MATS425, SPRING 2016

MIKKO PARVIAINEN
UNIVERSITY OF JYVÄSKYLÄ

Contents

1. Introduction

This two week intensive course is an introduction to the mathematical theory of optimal control. We mainly follow [Eva05] adding some material from [BCD97], [FS06], [Eva10].

We also study connections to partial differential equations and present several examples from both the practical and mathematical point of views.

Date: October 25, 2016.

Example 1.1. *Basic setup of an optimal control problem with (x,t) as a starting point We use the following terminology and notation*

$$
\begin{aligned}
&(x,t) \in \mathbb{R}^n \times [0,T], \textit{starting point}\\
&\alpha : [t,T] \to A, \textit{measurable}, A \textit{ a compact subset of } \mathbb{R}^n, \textit{ admissible control}\\
&\mathcal{A}, \textit{ the set of admissible controls}\\
&\textit{a trajectory } x(\cdot) \textit{is given by}\\
&\begin{cases} x'(s) = f(x(s),\alpha(s)), & s \in [t,T] \\ x(t) = x \end{cases} \textit{ dynamics given by } f\\
&f : \mathbb{R}^n \times A \to \mathbb{R}^n,\\
&r : \mathbb{R}^n \times A \to \mathbb{R}, \textit{ running cost},\\
&g : \mathbb{R}^n \to \mathbb{R}, \textit{ terminal/cost cost.}
\end{aligned}
$$

The value for this control problem is given by

$$u(x,t) = \inf_{\alpha \in \mathcal{A}} C_{x,t}(\alpha) := \inf_{\alpha \in \mathcal{A}} \int_t^T r(x(s),\alpha(s))\,ds + g(x(T)),$$

or we may want to maximize payoff

$$v(x,t) = \sup_{\alpha \in \mathcal{A}} P_{x,t}(\alpha) = \sup_{\alpha \in \mathcal{A}} \int_t^T r(x(s),\alpha(s))\,ds + g(x(T)).$$

A control $\tilde{\alpha} \in \mathcal{A}$ is optimal if for example in the first case $C_{x,t}(\tilde{\alpha}) = \inf_{\alpha \in \mathcal{A}} C_{x,t}(\alpha) = u(x,t)$.

Questions

- *How to find/contruct optimal control?*
- *Optimal control exist?*
- *Value of u?*

It turn out that u is a unique viscosity solution to Hamilton-Jacobi-Bellman first order PDE

$$\begin{cases} u_t + H(x,Du) = 0, & \textit{in } \mathbb{R}^n \times (t,T) \\ u(x,T) = g(x) & \textit{in } \mathbb{R}^n. \end{cases}$$

Thus finding u and α can sometimes transformed into a PDE question.

Remark 1.2. *First that the above makes sense, we need some assumptions on f, r, g, α. We return to assumptions later and first consider a few examples.*

Also t dependencies could be allowed above.

Example 1.3 (Production planning). *Factory produces n commodities.*

$$x(s) = \begin{pmatrix} x_1(s) \\ \vdots \\ x_n(s) \end{pmatrix} = \textit{inventory levels}, \quad \alpha(s) = \begin{pmatrix} \alpha_1(s) \\ \vdots \\ \alpha_n(s) \end{pmatrix} = \textit{production levels},$$

$$d = \begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix} = \textit{constant demand rates.}$$

The inventory develops according to

$$x'(s) = \alpha(s) - d.$$

Moreover there are constraints how much factory can produce each commodity

$$\alpha(s) \in A$$

where $A \subset \mathbb{R}^n_+$ is a compact set. Problem: choose α to minimize

$$C_{x,t}(\alpha) = \int_t^T r(x(s), \alpha(s))\, ds.$$

An example of r with $n = 1$

$$r(x,a) = \underbrace{x \cdot h}_{\textit{holding cost}} - \underbrace{rev \cdot d}_{\textit{revenue}} + \underbrace{c \cdot a}_{\textit{production cost}},$$

where $h, rev, c \geq 0$ are given constants. Remarks:

- *Case 1: we add a state constraint $x(s) \geq 0$.*
- *Revenue is constant and plays no role.*
- *We can fix the endpoint $x(T) = 0$ or let it be free.*
- *We can also put final cost.*
  Case 2: Alternative problem formulation. Drop state constraint, but add shortage penalty:

  $$r(x,a) = \underbrace{x^+ \cdot h}_{\textit{holding cost}} - \underbrace{rev \cdot d}_{\textit{revenue}} + \underbrace{c \cdot a}_{\textit{production cost}} + \underbrace{x^- \cdot pen}_{\textit{shortage penalty}},$$

  where $x^+ = \max(x, 0)$, $x^- = -\min(x, 0)$. Interpretation could now be that you pay when placing order, but then if we are unable to deliver there is a penalty. If $x < 0$, this can be though as how many undelivered orders are standing in line.

Example 1.4 (Investment planning).

$$x(t) = \textit{rate of revenue of our company at time } t,$$
$$\alpha(t) = \textit{reinvestment rate into the company} \in [0, 1].$$

Revenue develops according to

$$\begin{cases} x'(s) = k\alpha(s)x(s), \\ x(0) = x = \textit{initial revenue rate} > 0, \end{cases}$$

where k *=investment growth rate. Problem is to maximize profit (or dividends for example)*

$$P_{x,0}(\alpha) = \int_0^T x(s)(1-\alpha(s))\,ds.$$

- *Think of CEO whose bonus only depends profit over* $[0, T]$ *and knows that he leaves the company at time* T*.*
- *More realistic would be to have terminal condition too, e.g. share value at the end.*

The solution so called bang-bang control: first reinvestment everything and then nothing; we return later to this in Example 2.13.

Example 1.5 (Rocket railroad car)**.** *Consider a rocket railroad car with engines on both sides with the variables*

$$\begin{aligned} q(t) &= \textit{position at time } t \\ v(t) &= q'(t) = \textit{velocity at time } t \\ \alpha(t) &= \textit{control i.e. thrust (force) from the rockets}, \alpha(t) \in [-1, 1], \\ m &= \textit{mass of the rocket car} = 1. \end{aligned}$$

Q: How to operate the rockets to reach 0 *at the minimum amount of time and stop there? This problem can be formulated by using notation*

$$x(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} q(t) \\ v(t) \end{pmatrix}$$

as

$$x'(t) = \begin{pmatrix} q'(t) \\ v'(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} x(t) + \begin{pmatrix} 0 \\ \alpha(t) \end{pmatrix} =: f(x(t), \alpha(t)).$$

Above

$$mv'(t) = \textit{mass}\cdot\textit{acceleration}=\textit{force}=\alpha$$

is the law of motion. The total 'cost' to be minimized is

$$C_{x_0,0}(\alpha) := \int_0^\tau 1\,dt$$

where $\tau := \tau(\alpha, x_0)$*='first time* $q(t) = 0 = v(t)$*'. The value is*

$$u(x_0) := \inf_\alpha C_{x_0}(\alpha) = \inf_a \int_0^\tau 1\,dt.$$

Example 1.6 (Rocket car continued, heuristic solution)**.** *Guess that the optimal α only uses the extremal values -1 or 1. Some justification later. Recall*

$$x'(t) = f(x(t), \alpha(t)) = \begin{pmatrix} v \\ \alpha(t) \end{pmatrix} = \begin{pmatrix} x_2(t) \\ \alpha(t) \end{pmatrix}.$$

Whenever $\alpha \equiv 1$ we have

$$x'(t) = \begin{pmatrix} q'(t) \\ v'(t) \end{pmatrix} = \begin{pmatrix} v(t) \\ 1 \end{pmatrix}.$$

This implies

$$vv' = q' \Rightarrow \frac{1}{2}(v^2)' = q',$$

and furter

$$v^2(t) - v^2(0) = \int_0^t (v^2(s))'\, ds = 2\int_0^t q'(s)\, ds = 2q(t) - 2q(0).$$

Rearranging

$$v^2(t) = 2q(t) + \underbrace{(v^2(0) - 2q(0))}_{known} = 2q(t) + 2b. \tag{1.1}$$

Whenever $\alpha \equiv -1$

$$v^2(t) = -2q(t) + \underbrace{(v^2(0) + 2q(0))}_{known} = -2q(t) + 2\overline{b}. \tag{1.2}$$

Rewriting

$$\begin{cases} \frac{1}{2}v^2 - b = q, & \alpha = 1 \\ -\frac{1}{2}v^2 + \overline{b} = q, & \alpha = -1. \end{cases}$$

By drawing the images, we can now find optimal trajectories. For example, quite naturally, if we are on the positive side with zero initial speed we always use $\alpha = -1$ until we are close enough 0 (i.e. cross the parabola going through $(0,0)$) and then stop the car by using $\alpha = 1$. This is an example of a so called bang-bang control.

Suppose that $x(0) = (1,0)$. Then we first use $\alpha = -1$ and then $\alpha = 1$. To calculate how long it takes to reach the origin, solve

$$\int dv = \int -1\, dt \Rightarrow v(t) = -t + c$$

and since $v(0) = 0$ it follows that $c = 0$. Moreover,

$$\int dq = \int -t\, dt \Rightarrow q(t) = -\frac{1}{2}t^2 + \overline{c}.$$

Since $q(0) = \overline{c} = 1$, we have

$$q(t) = -\frac{1}{2}t^2 + 1.$$

By solving from (1.1) and (1.2) the time of the change in control, or arguing by symmetry that the change is in the midway, we see that the control changes at $t = 1$. Repeating a similar calculation or using symmetry, it takes

$$t = 2$$

to reach the origin.

More generally, in (q, v)-plane , the curve

$$v = \begin{cases} \sqrt{-2q}, & q < 0, \\ -\sqrt{2q}, & q \geq 0, \end{cases}$$

separates the two regions: in Region 1

$$v = \begin{cases} \geq \sqrt{-2q}, & q < 0, \\ > -\sqrt{2q}, & q \geq 0. \end{cases} \tag{1.3}$$

we start with $\alpha = -1$ whereas in the complement, Region 2, we start with $\alpha = 1$.

1.1. Classification of problems.

- Fixed finite time horizon, free endpoint, $\mathbb{R}^n \times [0, T]$(second case of 'Production planning', Example 1.3): find control that minimizes
$$\int_t^T r(x(s), \alpha(s))\, ds + g(x(T)).$$
- Fixed finite time horizon and control until exit from a cylindrical region $Q \times [0, T]$.
- Infinite time horizon: control until exit from a domain Q. Minimize the functional
$$\int_t^\tau r(x(s), \alpha(s))\, ds + g(x(\tau), \tau)\chi_{\{\tau<\infty\}}$$
where again $\tau := \tau(\alpha, x_0)$. Sometimes discounted payoffs are used
$$r(x, t) = e^{-\beta t}\tilde{r}(x, t),\ \ g(x, t) = e^{-\beta t}\tilde{g}(x, t)$$
to ensure finiteness of the payoffs in certain problems.
- Fixed endpoint problem (cf. the rocket car which had no discount):
$$\int_t^\tau r(x(s), \alpha(s))\, ds, \quad x(\tau) = x_1.$$
- Final endpoint constraint. Same as previous, but the target set can be some bigger set:
$$\int_t^\tau r(x(s), \alpha(s))\, ds + g(x(\tau(\alpha))), x(\tau) \in K.$$

- State constraint problem (first case of 'Production planning', Example 1.3). Only controls that guarantee
$$x(s) \in U \subset \mathbb{R}^n.$$

2. Finite time horizon

2.1. **Dynamic programming.** We continue by looking at finite time horizon.

Here are the **standing assumptions**. We assume that all the functions are continuous (in all variables) Lip wrt x, and bounded
$$\begin{aligned}
&|f(x,a)| \leq C, \ |r(x,a)| \leq C, \ |g(x)| \leq C \\
&|f(x,a) - f(y,a)| \leq C\,|x-y|\,, \ |r(x,a) - r(y,a)| \leq C\,|x-y|\,, \\
&|g(x) - g(y)| \leq C\,|x-y|\,.
\end{aligned}$$
Recall that $\alpha : [t,T] \to A$ is a measurable function with compact A, and that the dynamics is given by the ODE
$$\begin{cases} x'(s) = f(x(s), \alpha(s)) & t \leq s \leq T \\ x(t) = x. \end{cases}$$
It has a unique Lipschitz continuous solution $x(s)$ under these assumptions.

We are interested of the total costs and therefore define a value function
$$u(x,t) := \inf_{\alpha \in \mathcal{A}} C_{x,t}(\alpha) := \inf_{\alpha \in \mathcal{A}} \{ \int_t^T r(x(s), \alpha(s))\, ds + g(x(T)) \}.$$
We are going to show that u is a solution to a first order terminal value PDE of a type
$$\begin{cases} u_t + H(x, Du) = 0 & \text{in } \mathbb{R}^n \times (0,T) \\ u(x,T) = g(x) & \text{on } \mathbb{R}^n, \end{cases}$$
where
$$H(x,Du) := \inf_{a \in A} H(x,Du,a) := \inf_{a \in A} \{ f(x,a) \cdot Du(x,t) + r(x,a) \} \tag{2.4}$$
is the Hamiltonian, and $H(x, Du, a)$ is the control theory Hamiltonian. This is often called Hamilton-Jacobi-Bellman (HJB) equation, and can be seen as an infinitesimal version of so called dynamic programming principle (or Bellman's dynamic programming principle or optimality condition) that we will introduce next. It is the basis of the solution technique developed by Bellman in 1950s.

Lemma 2.1 (DPP/optimality condition)**.** *For each $h > 0$ small enough so that $t + h \leq T$, we have*
$$u(x,t) = \inf_{\alpha \in \mathcal{A}} \{ \int_t^{t+h} r(x(s), \alpha(s))\, ds + u(x(t+h), t+h) \},$$

where $x(\cdot)$ is the trajectory with the control α.

Idea is that we can think that we play optimally from time $t+h$ and thus obtain $u(x(t+h)t+h)$. Getting there costs $\int_t^{t+h} r(x(s),\alpha(s))\,ds$. For the proofs of the following theorems, see for example Section 10.3 in Evans: Partial differential equations.

The next theorem quarantees that we remain within the uniqueness theory in unbounded domain (See 10.2 in [Eva10]), and make sure that we take the initial values continuously.

Theorem 2.2 (Estimates for value functions)**.** *The above value function u satisfies*

$$\begin{aligned} &|u(x,t)| \le C, \\ &\left|u(x,t)-u(\hat{x},\hat{t})\right| \le C\,|x-\hat{x}| + C\left|t-\hat{t}\right|, \end{aligned}$$

for $x,\hat{x}\in\mathbb{R}^n$, $0\le t,\hat{t}\le T$.

Next we connect the value function to the PDE. The PDE then provides (existence, uniqueness, solvers available...) us the way to access the value function and the optimal control.

The heuristics is that the PDE is an infinitesimal version of DPP. Formally supposing u is a smooth value we can start from the DPP

$$0 = \inf_{\alpha\in\mathcal{A}}\{\int_t^{t+h} r(x(s),\alpha(s))\,ds + u(x(t+h),t+h) - u(x,t)\},$$

and assume that we are only using controls such that $\lim_{h\to 0+}\alpha(t+h)=a$. Then taking limit, changing order of lim and inf and dividing by h, we formally obtain

$$\begin{aligned} 0 &= \inf_{a\in A}\{r(x,a) + \frac{d}{dt}(u(x(t),t))\} \\ &= \inf_{a\in A}\{r(x,a) + Du(x(t),t)\cdot x'(t) + u_t(x(t),t))\} \\ &= \inf_{a\in A}\{r(x,a) + Du(x,t)\cdot f(x,a)\} + u_t(x,t)) \\ &= H(x,Du(x,t)) + u_t(x,t). \end{aligned}$$

The value should thus be related to the problem

$$\begin{cases} H(x,Du(x,t)) + u_t(x,t) = 0 \\ u(x,T) = g(x) \end{cases} \tag{2.5}$$

We recall the viscosity solutions, for more details see for example the lecture note of 'Viscosity theory'.

Definition 2.3 (Parabolic viscosity solution)**.** *A function $u:\mathbb{R}^n\times(0,T)\to(-\infty,\infty)$, $u\in C(\mathbb{R}^n\times(0,T))$ is a viscosity solution to (2.5) if u takes the* *final values continuously and whenever* $\varphi \in C^1(\mathbb{R}^n \times (0,T))$ *touches* u *at* $(x,t) \in \mathbb{R}^n \times (0,T)$ *from below, it holds that*

$$\varphi_t(x,t) + H(x, D\varphi) \leq 0,$$

(this is supersolution condition) as well as whenever $\varphi \in C^1(\mathbb{R}^n \times (0,T))$ *touches* u *at* $(x,t) \in \mathbb{R}^n \times (0,T)$ *from above, it holds that*

$$\varphi_t(x,t) + H(x, D\varphi) \geq 0,$$

(subsolution condition).

Remark 2.4.

- *If there is no such test function* φ*, then the condition is automatically satisfied.*
- *The name viscosity solution is for historical reasons. The origins of the above definition is in the method of vanishing viscosity, i.e. one adds a smoothing, viscosity term,* $+\varepsilon\Delta u$ *and let* $\varepsilon \to 0$*.*

Example 2.5. *Consider time independent case*

$$F(u_x) = -|u_x| + 1 = 0.$$

Then u *is a supersolution if a function* $\varphi \in C^1$ *touching from below at* x *satisfies*

$$-|\varphi_x(x)| + 1 \leq 0$$

and subsolution if from above

$$-|\varphi_x(x)| + 1 \geq 0.$$

Then

$$u(x) = 1 - |x|$$

is the unique viscosity solution to

$$\begin{cases} -|u_x| + 1 = 0 & x \in (-1,1) \\ u(\pm 1) = 0. \end{cases}$$

Theorem 2.6. *Let* u *be the value for the control problem in the above setup. Then* u *is a viscosity solution to*

$$\begin{cases} u_t(x,t) + H(x, Du) = 0 & \text{in } \mathbb{R}^n \times (0,T) \\ u(x,T) = g(x). \end{cases}$$

Proof. Let u be a value. Thriving for a contradiction, assume that the subsolution property does not hold i.e. there is $\varphi \in C^1(\mathbb{R}^n \times (0,T))$ touching u at (x,t) from above

$$\varphi_t(x,t) + H(x, D\varphi(x,t)) = \varphi(x,t)_t + \inf_{a \in A}\{f(x,a) \cdot D\varphi(x,t) + r(x,a)\} < 0.$$

Then by continuity there is $a \in A$ and $\theta > 0$ such that

$$\varphi_t(y,s) + f(y,a) \cdot D\varphi(y,s) + r(y,a) < -\theta \tag{2.6}$$

for all points (y, s) sufficiently close to (x, t).

Choose a fixed control $\alpha(s) \equiv a$ and the corresponding trajectory

$$\begin{cases} x'(s) = f(x(s), a), & t \le s \le T, \\ x(t) = x. \end{cases}$$

By continuity of the trajectory and (2.6), it holds for small enough h, $t \le s \le t + h$ that

$$\varphi_t(x(s), s) + f(x(s), a) \cdot D\varphi(x(s), s) + r(x(s), a) < -\theta \tag{2.7}$$

Thus, since φ touches u from above

$$\begin{aligned} &u(x(t+h), t+h) - u(x(t), t) \\ &\quad \le \varphi(x(t+h), t+h) - \varphi(x(t), t) \\ &\quad = \int_t^{t+h} \frac{d}{ds} \varphi(x(s), s)\, ds \\ &\quad = \int_t^{t+h} D\varphi(x(s), s) \cdot \underbrace{x'(s)}_{= f(x(s), a)} + \varphi_t(x(s), s)\, ds. \end{aligned}$$

On the other hand, by the DPP

$$\begin{aligned} u(x, t) &= \inf_{\alpha \in \mathcal{A}} \{ \int_t^{t+h} r(x(s), \alpha(s))\, ds + u(x(t+h), t+h) \} \\ &\le \int_t^{t+h} r(x(s), a)\, ds + u(x(t+h), t+h). \end{aligned}$$

Combining the previous two inequalities, we end up with

$$\begin{aligned} &u(x(t+h), t+h) - (\int_t^{t+h} r(x(s), a)\, ds + u(x(t+h), t+h)) \\ &\quad \le u(x(t+h), t+h) - u(x(t), t) \\ &\quad \le \int_t^{t+h} D\varphi(x(s), s) \cdot f(x(s), a) + \varphi_t(x(s), s)\, ds \end{aligned}$$

i.e.

$$0 \le \int_t^{t+h} r(x(s), a) + D\varphi(x(s), s) \cdot f(x(s), a) + \varphi_t(x(s), s)\, ds \overset{(2.7)}{<} -h\theta,$$

a contradiction.

To prove that supersolution property holds for u, let φ touch u at (x, t) from below, and suppose that supersolution property does not hold. By continuity,

$$\varphi(x(s), s)_t + f(x(s), \alpha(s)) \cdot D\varphi(x(s), s) + r(x(s), \alpha(s)) > \theta. \tag{2.8}$$

for any $\alpha \in \mathcal{A}$ and s close enough t.

Thus for any control $\alpha \in \mathcal{A}$, not necessarily constant, and the corresponding trajectory $x(s)$. Further, since φ touches u from below

$$\begin{aligned} &u(x(t+h), t+h) - u(x(t), t) \\ &\quad \geq \varphi(x(t+h), t+h) - \varphi(x(t), t) \\ &\quad = \int_t^{t+h} \frac{d}{ds}\varphi(x(s), s)\, ds \\ &\quad = \int_t^{t+h} D\varphi(x(s), s) \cdot f(x(s), \alpha(s)) + \varphi_t(x(s), s)\, ds, \end{aligned}$$

for any strategy. Further by the DPP and definition of inf, for $\theta/2 > 0$, there is $\alpha \in \mathcal{A}$ such that

$$u(x, t) \geq \int_t^{t+h} r(x(s), \alpha(s))\, ds + u(x(t+h), t+h) - h\theta/2$$

and combining the previous two inequalities

$$\begin{aligned} &u(x(t+h), t+h) - (\int_t^{t+h} r(x(s), \alpha(s))\, ds + u(x(t+h), t+h) - h\theta/2) \\ &\geq \int_t^{t+h} D\varphi(x(s), s) \cdot f(x(s), \alpha(s)) + \varphi_t(x(s), s)\, ds. \end{aligned}$$

implying

$$h\theta/2 \geq \int_t^{t+h} r(x(s), \alpha(s)) + D\varphi(x(s), s) \cdot f(x(s), \alpha(s)) + \varphi_t(x(s), s)\, ds \overset{(2.8)}{>} h\theta,$$

a contradiction. □

2.1.1. *Using dynamic programming in designing optimal controls.*

(1) Find unique solution to the Hamilton-Jacobi-Bellman equation

$$\begin{cases} u_t(x, t) = H(x, Du) & \text{in } \mathbb{R}^n \times (0, T) \\ u(x, T) = g(x), & x \in \mathbb{R}^n. \end{cases}$$

According to Theorem 2.6, this is the value function.

(2) Solve *feedback control* by setting $\alpha(x, s) = \tilde{a}$ where

$$r(x, \tilde{a}) + Du(x, s) \cdot f(x, \tilde{a}) = \inf_{a \in A}\{r(x, a) + Du(x, s) \cdot f(x, a)\}.$$

(3) Solve ODE

$$\begin{cases} \tilde{x}'(s) = f(\tilde{x}(s), \alpha(\tilde{x}(s), s)) \\ \tilde{x}(t) = x \end{cases} \tag{2.9}$$

Define

$$\tilde{\alpha}(s) := \alpha(\tilde{x}(s), s)$$

a feedback control so that

$$u_t(\tilde{x}(s),s) + r(\tilde{x}(s),\tilde{\alpha}(s)) + Du(\tilde{x}(s),s)\cdot f(\tilde{x}(s),\tilde{\alpha}(s)) = 0. \tag{2.10}$$

Remark 2.7. *There are serious problems in the above procedure. There might be multiple solutions α, measurable selection theorem might be needed. Solving $x'(s) = f(\tilde{x}(s),\alpha(\tilde{x}(s),s))$ requires reqularity for α and does not always work, consider Example 2.9 in Reg II.*

Nonetheless in case all the problems indicated above can be solved, then the produced control $\tilde{\alpha}$ is optimal: Since the value and a solution to PDE coincide

$$\begin{aligned}
C_{x,t}(\tilde{\alpha}) &= \int_t^T r(\tilde{x}(s),s)\,ds + g(\tilde{x}(T)) \\
&\overset{(2.10)}{=} \int_t^T -u_t(\tilde{x}(s),s) - Du(\tilde{x}(s),s)\cdot f(\tilde{x}(s),\tilde{\alpha}(s))\,ds + g(\tilde{x}(T)) \\
&\overset{(2.9)}{=} \int_t^T -u_t(\tilde{x}(s),s) - Du(\tilde{x}(s),s)\cdot \tilde{x}'(s)\,ds + g(\tilde{x}(T)) \\
&= -\int_t^T \frac{d}{ds}u(\tilde{x}(s),s)\,ds + g(\tilde{x}(T)) \\
&= \underbrace{-u(\tilde{x}(T),T)}_{-g(\tilde{x}(T))} + u(\tilde{x}(t),t) + g(\tilde{x}(T)) \\
&= u(x,t) = \inf_{\alpha\in\mathcal{A}} C_{x,t}(\alpha).
\end{aligned}$$

Example 2.8. *Consider a problem in $\mathbb{R}\times[0,T]$ with controls $\alpha : [t,T] \to [-1,1]$ and dynamics*

$$\begin{cases} x'(s) = f(x(s),\alpha(s)) = \alpha(s), & t \le s \le T \\ x(t) = x, \end{cases}$$

and minimize (no running payoff)

$$C_{x,t}(\alpha) = g(x(T)) := e^{-x(T)^2}.$$

We guess that the optimal control is

$$\tilde{\alpha} = \begin{cases} 1, & x \ge 0 \\ -1, & x < 0. \end{cases}$$

(At $x = 0$, we are free to choose $\tilde{\alpha} = 1$ or $\tilde{\alpha} = -1$. Obviously optimal controls are not unique.) In any case

$$\tilde{x}(T) = \begin{cases} \int_t^T 1\,ds + x, & x \ge 0, \\ \int_t^T -1\,ds + x, & x < 0 \end{cases} = \begin{cases} T - t + x, & x \ge 0, \\ t - T + x, & x < 0 \end{cases} \tag{2.11}$$

and

$$u(x,t) = g(\tilde{x}(T)) = \begin{cases} e^{-(T-t+x)^2}, & x \geq 0, \\ e^{-(t-T+x)^2}, & x < 0. \end{cases}$$

Observe that

$$\lim_{y \to 0+} u_x(y,t) = -2(T-t)e^{-(T-t)^2} \neq -2(t-T)e^{-(t-T)^2} = \lim_{y \to 0-} u_x(y,t),$$

so that $u \notin C^1$ *even if the data is smooth.*

To different direction, suppose that we are first able to solve u *from the corresponding HJB equation* $u_t - |u_x| = 0$*. Assume* $x \neq 0$ *and solve* $\tilde{a}$ *from*

$$u_x(x,t)\tilde{a} = \inf_{a \in A}\{u_x(x,t)a\} = -|u_x|$$

i.e. $\tilde{a} = -\operatorname{sgn}(u_x(x,t))$*. Since*

$$u_x(x,t) = \begin{cases} -2(T-t+x)e^{-(T-t+x)^2}, & x > 0 \\ -2(t-T+x)e^{-(t-T+x)^2}, & x < 0, \end{cases}$$

we have

$$\alpha(x,t) = \tilde{a} = \begin{cases} 1, & x > 0 \\ -1, & x < 0. \end{cases}$$

Then we solve the full control by putting in trajectory

$$\begin{cases} \tilde{x}'(s) = \alpha(\tilde{x}(s), s), \\ \tilde{x}(t) = x, \end{cases}$$

Thus the optimal control related to the point (x,t) *is*

$$\tilde{\alpha}(s) = \tilde{\alpha}(\tilde{x}(s), s) = \begin{cases} 1, & x > 0, \\ -1, & x < 0. \end{cases}$$

This already indicates that there is problem of interpreting $Du(0)$*.*

Example 2.9 (Dynamics with three velocities)**.** *Let us consider the problem*

$$\begin{cases} x'(s) = \alpha(s), & 0 \leq t \leq s \leq 1 \\ x(t) = x, \end{cases}$$

where the control parameter takes values in

$$A = \{-1, 0, 1\}.$$

Our problem is to minimize

$$u(x,t) := C_{x,t}(\tilde{\alpha}) = \int_t^1 |x(s)| \, ds.$$

To this end, we consider three regions

$$\begin{aligned} I &= \{(x,t) : x \leq t-1, 0 \leq t \leq 1\}, \\ II &= \{(x,t) : t-1 < x < 1-t, 0 \leq t \leq 1\}, \\ III &= \{(x,t) : x \geq -t+1, 0 \leq t \leq 1\}. \end{aligned}$$

The value is

$$\begin{aligned} u(x,t) := C_{x,t}(\tilde{\alpha}) &= \begin{cases} \frac{1}{2}(-x+(-x-(1-t)))(1-t), & I \\ \frac{1}{2}x^2, & II \\ \frac{1}{2}(x+(x-(1-t)))(1-t), & III \end{cases} \\ &= \begin{cases} \frac{1}{2}(-2x-1+t)(1-t), & I \\ \frac{1}{2}x^2, & II \\ \frac{1}{2}(2x-1+t)(1-t), & III. \end{cases} \end{aligned} \tag{2.12}$$

Then the Hamilton-Jacobi-Bellman equation is

$$\begin{aligned} 0 &= u_t + \inf_{a \in A} \{f(x,a) \cdot u_x(x,t) + |x|\} \\ &= u_t + \inf_{a \in A} \{a \cdot u_x(x,t) + |x|\}. \end{aligned}$$

Then in Region II, it holds

$$\begin{aligned} &0 = u_t, \quad u_x = x, \\ &\inf_{a \in A} \{a \cdot u_x(x,t) + |x|\} = \inf_{a \in A} \{ax + |x|\} = -|x| + |x| = 0. \end{aligned}$$

in Region I

$$\begin{aligned} &u_t = \frac{1}{2}(1-t) - \frac{1}{2}(-2x-1+t) = 1-t+x, \quad u_x = -(1-t) \\ &\inf_{a \in A} \{a \cdot u_x(x,t) + |x|\} = \inf_{a \in A} \{-a(1-t) + |x|\} = -(1-t+x) \end{aligned}$$

so that HJB equation holds; the boundaries will have to be checked separately.

Example 2.10. *Consider the previous example with $A = \{-1, 1\}$. Then it is an exercise to show that there is no optimal control in Region II.*

However, value still exists and is the same as in (2.12), and it still satisfies the HJB.

2.2. **Pontryagin max principle for finite time horizon, free end-point.**

Next we encounter a necessary conditions for optimality (named after a soviet mathematician Lev Pontryagin 1908-1988) formulated in 50's.

Consider again the control problem

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

and try to find a control $\tilde{\alpha}$ such that (this time we maximize to get what is called max principle instead of minimum principle)

$$P_{x,t}(\tilde{\alpha}) = \sup_{\alpha\in\mathcal{A}} P_{x,t}(\alpha) = \sup_{\alpha\in\mathcal{A}} \int_0^T r(x(s),\alpha(s))\,ds + g(x(T)) =: u(x,0)$$

and with **the additional assumptions** $f(\cdot,a), r(\cdot,a), g \in C^1(\mathbb{R}^n)$.

Recall

$$H(x,Du) := \sup_{a\in A} H(x,Du,a) := \sup_{a\in A}(f(x,a)\cdot Du(x,t) + r(x,a))$$

is the Hamiltonian, and $H(x,Du,a)$ is the control theory Hamiltonian. After the calculation in Section 2.1.1, the following is rather natural.

Theorem 2.11 (Pontryagin max principle)**.** *Let $\tilde{\alpha}$ be optimal for the above problem and $\tilde{x}$ the corresponding trajectory. Then there exists a function (called costate or adjoint variable) $\tilde{p}(t) : [0,T] \to \mathbb{R}^n$ such that for a.e. $0 \le t \le T$*

$$\begin{aligned} \tilde{x}'(t) &= D_p H(\tilde{x}(t),\tilde{p}(t),\tilde{\alpha}(t)) \quad \textit{(ODE)} \\ \tilde{p}'(t) &= -D_x H(\tilde{x}(t),\tilde{p}(t),\tilde{\alpha}(t)) \quad \textit{(adjoint equations)=(ADJ)} \\ H(\tilde{x}(t),\tilde{p}(t),\tilde{\alpha}(t)) &= \sup_{a\in A} H(\tilde{x}(t),\tilde{p}(t),a) \quad \textit{(maximization princ)=(M)}, \\ \tilde{p}(T) &= Dg(\tilde{x}(T)) \quad \textit{(terminal/transversality condition)}. \end{aligned}$$

Recalling the control theory Hamiltonian we have

$$\tilde{x}'(t) = D_p H(\tilde{x}(t),\tilde{p}(t),\tilde{\alpha}(t)) = f(\tilde{x}(t),\tilde{\alpha}(t))$$

for (ODE), so this is just the usual dynamics.

Remark 2.12. *Observe that Pontryagin max principle might very well give false control candidates.*

It is rather a design tool not characterization: it is of course desirable whenever it only gives a small set of controls or only a single control.

Nonetheless, in principle, if we can make up a reasonable candidate for optimal control and thus derive a candidate as a value function, then we can check whether this is the unique viscosity solution to the corresponding PDE.

Next we see how to use Pontryagin max principle.

Example 2.13 (Investment planning)**.** *Recall the model:*

$$\begin{aligned} x(t) &= \textit{profit rate}, \\ \alpha(t) &= \textit{reinvestment rate} \in [0,1]. \end{aligned}$$

Profit develops according to

$$\begin{cases} x'(s) = \alpha(s)x(s) := f(x(s),\alpha(s)), \\ x(0) = x = \textit{initial profit rate} > 0. \end{cases}$$

. Problem is to maximize the profit

$$P_{x,0}(\alpha) = \int_0^T r(x(s),\alpha(s))\,ds + 0 = \int_0^T x(s)(1-\alpha(s))\,ds.$$

Thus the control theory Hamiltonian is

$$H(x,p,a) = f(x,a)p + r(x,a) = axp + x(1-a) = x(a(p-1)+1).$$

and equations from Pontryagin max principle

$$\begin{aligned} \tilde{x}'(s) &= \tilde{\alpha}(s)\tilde{x}(s) \quad \textit{(ODE)} \\ \tilde{p}'(s) &= -\tilde{\alpha}(s)(\tilde{p}(s)-1)-1, \quad \textit{(adjoint equations)} \\ \tilde{x}(s)(\tilde{\alpha}(s)(\tilde{p}(s)-1)+1) &= \sup_{a\in A}\{\tilde{x}(s)(a(\tilde{p}(s)-1)+1)\} \quad \textit{(maximization princ)}, \\ \tilde{p}(T) &= 0 \quad \textit{(terminal/transversality condition)}. \end{aligned}$$

From the maximization principle, since $\tilde{x} > 0$ by (ODE), we have

$$\alpha(s) = \begin{cases} 1, & \tilde{p}(s) > 1 \\ 0, & \tilde{p}(s) \le 1. \end{cases}$$

We are looking for a continuous costate function, and thus by terminal condition

$$\tilde{p}(s) \le 1 \text{ for } s \text{ close to } t.$$

Thus by (ADJ) for $s \in [T-1, T]$ (assume $T > 1$) we have

$$\tilde{p}(s) = T - s,$$

and by (ADJ) p' remains a.e. strictly negative and by (M) the control swithces at $T-1$, and there won't be further changes in the control. The optimal control is a.e. thus

$$\tilde{\alpha}(s) = \begin{cases} 1, & 0 \le s < T-1 \\ 0, & T-1 \le s \le T, \end{cases}$$

$$\tilde{x}(s) = \begin{cases} xe^{s}, & 0 \le s < T-1, \\ xe^{T-1}, & T-1 \le s \le T. \end{cases}$$

The payoff is

$$P_{x,0}(\tilde{\alpha}) = \int_{T-1}^{T} xe^{T-1}\,ds = xe^{T-1}.$$

22.1.2016

Theorem 2.14 (Costate and gradient)**.** *Assume that $\alpha(s) = \alpha_{x,t}(s)$ and $x(s) = x_{x,t}(s)$ solve the control problem at x, t. If the value function $u \in C^2$, then the costate p in Pontryagin is given by*

$$p(s) = Du(x(s),s), \quad t \le s \le T.$$

Proof. Let $p(s) := Du(x(s), s)$, then we claim that p satisfies (ADJ) and (M). With the optimal control it holds that

$$u_t(x(s), s) + r(x(s), \alpha(s)) + f(x(s), \alpha(s)) \cdot Du(x(s), s) = 0 \tag{2.13}$$

for a.e. $t \le s \le T$. Indeed, first observe that $x_{x_{x,t}(s),s}(\tilde{s}) = x_{x,t}(\tilde{s})$ for $t \le s \le \tilde{s} \le T$ when using the same control (or part of it) in both the cases. Thus a.e. s

$$\begin{aligned} \frac{d}{ds} u(x_{x,t}(s), s) &= \frac{d}{ds}\big(g(x_{x,t}(T)) + \int_s^T r(x_{x,t}(\tilde{s}), \alpha(\tilde{s}))\, d\tilde{s}\big) \\ &= -r(x_{x,t}(s), \alpha(s)). \end{aligned}$$

and

$$\begin{aligned} \frac{d}{ds} u(x_{x,t}(s), s) &= Du(x_{x,t}(s), s) \cdot x'_{x,t}(s) + u_t(x_{x,t}(s), s) \\ &= Du(x_{x,t}(s), s) \cdot f(x_{x,t}(s), \alpha_{x,s}(s)) + u_t(x_{x,t}(s), s) \end{aligned}$$

implying (2.13). Fix s where (2.13) holds, and define

$$h(y) := u_t(y, s) + r(y, \alpha(s)) + f(y, \alpha(s)) \cdot Du(y, s) \overset{\text{dropped max}}{\le} 0.$$

Moreover, $h(x(s)) = 0$ and thus h obtains its max at $y = x(s)$ so that

$$\begin{aligned} 0 = h_{x_i}(x(s)) &= u_{tx_i}(x(s), s) + r_{x_i}(x(s), \alpha(s)) \\ &\quad + f_{x_i}(x(s), \alpha(s)) \cdot Du(x(s), s) + f(x(s), \alpha(s)) \cdot Du_{x_i}(x(s), s). \end{aligned}$$

On the other hand, by definition of p and the above equation

$$\begin{aligned} p_i'(s) &= \frac{d}{ds} u_{x_i}(x(s), s) \\ &= u_{x_i t}(x(s), s) + Du_{x_i}(x(s), s) \cdot x'(s) \\ &= u_{x_i t}(x(s), s) + Du_{x_i}(x(s), s) \cdot f(x(s), \alpha(s)) \\ &= -r_{x_i}(x(s), \alpha(s)) - f_{x_i}(x(s), \alpha(s)) \cdot Du(x(s), s) \\ &\quad - f(x(s), \alpha(s)) \cdot Du_{x_i}(x(s), s) + Du_{x_i}(x(s), s) \cdot f(x(s), \alpha(s)) \\ &= -r_{x_i}(x(s), \alpha(s)) - f_{x_i}(x(s), \alpha(s)) \cdot Du(x(s), s). \end{aligned}$$

This is (ADJ).

Then comparing

$$u_t(x(s), s) + \sup_{a \in A} \{ r(x(s), s) + f(x(s), a) \cdot Du(x(s), s) \} = 0$$

with (2.13), we obtain (M). □

2.3. **Viscosity solutions and Pontryagin max principle.** Next we make few comments on how Barron and Jensen, 1986, TAMS used ideas of Theorem 2.14 in a rigorous proof using viscosity solutions.

We continue under the assumptions of the Pontryagin max principle. Fix the optimal control

$$\tilde{\alpha}_{x,t} : [t, T] \to A$$

related to the starting point (x, t) and

$$\begin{cases} \tilde{x}'_{x,t}(s) = f(\tilde{x}_{x,t}(s), \tilde{\alpha}_{x,t}(s) \\ \tilde{x}_{x,t}(t) = x. \end{cases}$$

Then define $w : \mathbb{R}^n \times [t, T] \to \mathbb{R}$ by

$$w(\xi, \tau) = g(x_{\xi,\tau}(T)) + \int_{\tau}^{T} r(x_{\xi,\tau}(s), \tilde{\alpha}_{x,t}(s)) \, ds,$$

where $t \le \tau \le s \le T$ and

$$\begin{cases} x'_{\xi,\tau}(s) = f(x_{\xi,\tau}(s), \tilde{\alpha}_{x,t}(s)) \\ x_{\xi,\tau}(\tau) = \xi. \end{cases}$$

Idea is now to show that the 'value' w with the frozen control has enough regularity (which we take for granted) and this gives us (M) with $\tilde{p}(s) := Dw(\tilde{x}_{x,t}(s), s)$.

Theorem 2.15. *It holds that*

$$\begin{aligned} &Dw(\tilde{x}_{x,t}(s), s) \cdot f(\tilde{x}_{x,t}(s), \tilde{\alpha}_{x,t}(s)) + r(\tilde{x}_{x,t}(s), \tilde{\alpha}_{x,t}(s)) \\ &\quad = sup_{a \in A}\{Dw(\tilde{x}_{x,t}(s), s) \cdot f(\tilde{x}_{x,t}(s), a) + r(\tilde{x}_{x,t}(s), a)\}, \end{aligned}$$

for almost every $s \in [t, T]$.

Proof. It holds that

$$\begin{aligned} w(y, T) &= g(y) \\ w(x, t) &= u(x, t) \\ w(\tilde{x}_{x,t}(s), s) &= u(\tilde{x}_{x,t}(s), s) \end{aligned}$$

and

$$u(\xi, \tau) \ge w(\xi, \tau) \quad \text{for } (\xi, \tau) \in \mathbb{R}^n \times [t, T]$$

since u is the largest payoff at each point. That is, on each point on the trajectory $\tilde{x}_{x,t}(s)$ the function w touches u from below. Since u is known to be viscosity supersolution it holds, assuming w is differentiable on a point $(\tilde{x}_{x,t}(s), s)$, on the trajectory, that

$$w_t(\tilde{x}_{x,t}(s), s) + \sup_{a \in A}\{Dw(\tilde{x}_{x,t}(s), s) \cdot f(\tilde{x}_{x,t}(s), a) + r(\tilde{x}_{x,t}(s), a)\} \le 0.$$

The next lemma proves the reverse inequality. □

Lemma 2.16. *At almost every* $(s, x_{\xi,\tau}(s))$, $t \le \tau \le s \le T$, *we have*

$$w_t(x_{\xi,\tau}(s), s) + Dw(x_{\xi,\tau}(s), s) \cdot f(x_{\xi,\tau}(s), \tilde{\alpha}_{x,t}(s)) + r(x_{\xi,\tau}(s), \tilde{\alpha}_{x,t}(s)) = 0.$$

Proof. Idea is the same as in Theorem 2.14. Again assume enough regularity for w. By definition of w it holds that

$$\begin{aligned} \frac{d}{ds} w(x_{\xi,\tau}(s), s) &= \frac{d}{ds}\Big(g(x_{\xi,\tau}(T)) + \int_s^T r(x_{\xi,\tau}(\tilde{s}), \tilde{\alpha}_{x,t}(\tilde{s}))\, d\tilde{s} \Big) \\ &= -r(x_{\xi,\tau}(s), \tilde{\alpha}_{x,t}(s)). \end{aligned}$$

where we also used the fact $x_{x_{\xi,\tau}(s),s}(\tilde{s}) = x_{\xi,\tau}(\tilde{s})$ for $\tilde{s} \ge s$. On the other hand, by the chain rule

$$\begin{aligned} \frac{d}{ds} w(x_{\xi,\tau}(s), s) &= Dw(x_{\xi,\tau}(s), s) \cdot x'_{\xi,\tau}(s) + w_t(x_{\xi,\tau}(s), s) \\ &= Dw(x_{\xi,\tau}(s), s) \cdot f(x_{\xi,\tau}(s), \tilde{\alpha}_{x,t}(s)) + w_t(x_{\xi,\tau}(s), s). \end{aligned}$$ □

2.4. **Pontryagin max principle and PDEs.** Next we look how to come close to (ODE) and (ADJ)

$$\begin{aligned} \tilde{x}'(t) &= D_p H(\tilde{x}(t), \tilde{p}(t), \tilde{\alpha}(t)) \quad \text{(ODE)} \\ \tilde{p}'(t) &= -D_x H(\tilde{x}(t), \tilde{p}(t), \tilde{\alpha}(t)) \quad \text{(ADJ)}. \end{aligned}$$

from the point of view of the PDE theory. They arise through the method of characteristics, i.e. we try to find such trajectories that we can calculate the solution for the Hamilton-Jacobi PDE

$$\begin{cases} u_t + H(x, Du(x,t)) = 0 & \text{in } \mathbb{R}^n \times (0,T) \\ u(x,0) = g(x) & \text{in } \mathbb{R}^n. \end{cases}$$

We assume that u is *smooth* and denote

$$\begin{aligned} x(s) &:= \begin{pmatrix} x_1(s) \\ \vdots \\ x_n(s) \end{pmatrix}, \\ p(s) &:= D_x u(x(s), s) = \begin{pmatrix} p_1(s) \\ \vdots \\ p_n(s) \end{pmatrix} = \begin{pmatrix} u_{x_1}(x(s), s) \\ \vdots \\ u_{x_n}(x(s), s) \end{pmatrix}. \end{aligned} \tag{2.14}$$

We differentiate Hamilton-Jacobi equation wrt x_k to get

$$u_{tx_k}(x,t) + H_{x_k}(x, Du(x,t)) + \sum_i H_{p_i}(x, Du) u_{x_k x_i}(x,t) = 0. \tag{2.15}$$

Using the notation introduced above

$$\begin{aligned}
p_k'(t) &= \frac{d}{dt} u_{x_k}(x(t),t) \\
&= u_{tx_k}(x(t),t) + \sum_i u_{x_k x_i}(x(t),t) x_i'(t) \\
&\overset{(2.15)}{=} -H_{x_k}(x(t), Du(x(t),t)) - \sum_i H_{p_i}(x(t), Du(x(t),t)) u_{x_k x_i}(x(t),t) \\
&\qquad\qquad + \sum_i u_{x_k x_i}(x(t),t) x_i'(t) \\
&= -H_{x_k}(x(t), Du(x(t),t)) + \sum_i (-H_{p_i}(x(t), Du(x(t),t)) + x_i'(t)) u_{x_k x_i}(x(t),t) \\
&= -H_{x_k}(x(t), p(t)) + \sum_i (-H_{p_i}(x(t), p(t)) + x_i'(t)) u_{x_k x_i}(x(t),t).
\end{aligned}$$

We want to simplify, and to this end get rid of second derivatives, by choosing

$$x_i'(t) = H_{p_i}(x(t), p(t)).$$

Then the above further yields

$$p_k'(t) = -H_{x_k}(x(t), p(t)).$$

These are Hamilton's equations. If we can solve these with suitable initial conditions $x(0) = x_0, p(0) := p_0 := Dg(x_0)$, then when can calculate the solution along the trajectories:

$$\begin{aligned}
u(x(t),t) &= \int_0^t \frac{d}{ds} u(x(s),s)\, ds + u(x(0),0) \\
&= \int_0^t \big(Du(x(s),s) \cdot x'(s) + u_t(x(s),s)\big)\, ds + u(x(0),0) \\
&\overset{\text{(ADJ),(ODE),(2.14)}}{=} \int_0^t \big(p(s) \cdot D_p H(x(s),p(s)) - H(x(s),p(s))\big)\, ds + u(x(0),0).
\end{aligned}$$

Example 2.17. *Consider*

$$\begin{cases} u_t + |u_x| = 0 \\ u(x,0) = g(x) = e^{-x^2} \end{cases}.$$

Then away from $p = 0$ *we use method of characteristic,*

$$\begin{aligned}
H(x,p) = |p|\,, H_p = \operatorname{sgn} p, H_x = 0 \\
p(t) = p_0 = Dg(x_0) = -2x_0 e^{-x_0^2}, \\
x'(t) = H_p(x(t),p(t)) = \operatorname{sgn}(-2x_0 e^{-x_0^2}) = -\operatorname{sgn}(x_0)
\end{aligned}$$

so that

$$\begin{aligned} u(x(t),t) &= \int_0^t \big(p(s)\cdot D_pH(x(s),p(s)) - H(x(s),p(s))\big)\,ds + u(x(0),0) \\ &= \int_0^t \big(-p_0\,\mathrm{sgn}(x_0) - |p_0|\big)\,ds + g(x_0) \\ &= t(2\,|x_0|\,e^{-x_0^2} - 2\,|x_0|\,e^{-x_0^2}) + e^{-x_0^2} \\ &= e^{-x_0^2}. \end{aligned}$$

On the other hand we might consider through Pontryagin the 1D minimization problem we already solved before (related to $u_t - |u_x| = 0$)

$$g(x(T)) = e^{-x(T)^2}$$

under

$$\begin{cases} x'(s) = \alpha(s), \\ x(t) = x. \end{cases}$$

This time we have Pontryagin minimization principle

$$\begin{aligned} p'(t) &= 0, \\ p(T) &= Dg(x(T)) = -2x(T)e^{-x(T)^2}, \\ p(s)\cdot\alpha(s) &= \min_{a\in[-1,1]} p(s)a. \end{aligned}$$

If $X(T) > 0$ then $p(t) = p_0 < 0$ and by (M), $\alpha(s) \equiv 1$. If $X(T) < 0$, then $\alpha(s) \equiv -1$.

3. Heuristic discussions and warnings on Pontryagin

The following 'argument' leaves a lot to be desired, but might still provide some insight.

Consider a problem with exit time constrained to exit on a set $\partial\mathcal{T}$, (like the rocket car example, we return to this in Section 5). Let $\tau = \tau(x,\alpha)$ be the first hitting time of the set. Add the trajectory as a penalty term and integrate by parts

$$\begin{aligned} &\int_0^\tau r(x(t),\alpha(t)) + p(t)\cdot(f(x(t),\alpha(t)) - x'(t))\,dt + g(x(\tau)) \\ &\quad= \int_0^\tau r(x(t),\alpha(t)) + p(t)\cdot f(x(t),\alpha(t))\,dt + g(x(\tau)) \\ &\qquad - p(\tau)\cdot x(\tau) + p(0)\cdot x(0) + \int_0^\tau p'(t)x(t)\,dt. \end{aligned}$$

Max of quantity

$$-p(\tau)\cdot x(\tau) + g(x(\tau))$$

where $x(\tau) \in \partial\mathcal{T}$, under the condition $h(x(\tau)) = 0$(where h is the equation of the boundary, assume smooth, $Dh(x(\tau)) \neq 0$) is according to Lagrange multipliers (see Evans' lecture note) requires

$$-p(\tau) + Dg(x(\tau)) = cDh(x(\tau))$$

i.e.

$$(-p(\tau) + Dg(x(\tau))) \cdot \bar{t} = 0$$

where $\bar{t}$ is a vector in a tangent plane to $\partial\mathcal{T}$ at $x(\tau)$. This gives the transversality condition; also in the fixed endpoint or free endpoint problems.

Remark 3.1. *Also it should be noted that the version of the Pontryagin stated is not exactly accurate. There are so called abnormal problems for which the control Hamiltonian reads as $H(x, p, a) = f \cdot p + 0 \cdot r$ and for which Pontryagin is not useful.*

Following example is due to Arytyonov, Differential Equations, 2010. Consider the rocket car example, take the essentially unique optimal control α steering us from $x_0 \in \mathbb{R}^2$ to 0 in optimal time T. Choose x_0 so that the control has a unique switching time $\tau < T$ for the control. Then consider the following control problem: minimize

$$\int_0^T \alpha(t)\sqrt{|\tau - s|}\, ds$$

with fixed starting x_0 and end point 0. This is fixed time, fixed starting and fixed endpoint problem. The above α is (essentially) the only admissible control and thus optimal. Try to follow Pontryagin with

$$\begin{aligned} H(x, p, a, s) &= f(x, a) \cdot p + r(x, a, s) \\ &= a\sqrt{|\tau - s|} \\ &= x_2 p_1 + p_2 a + a\sqrt{|\tau - s|}. \end{aligned}$$

so that

$$\begin{aligned} &\begin{pmatrix} p_1'(s) \\ p_2'(s) \end{pmatrix} = -\begin{pmatrix} 0 \\ p_1(s) \end{pmatrix} \ \textit{(ADJ)}, \\ &x_2(s)p_1(s) + p_2(s)\alpha(s) + \alpha(s)\sqrt{|\tau - s|} \\ &\quad = \min_{a \in [-1,1]} \{x_2(s)p_1(s) + p_2(s)a + a\sqrt{|\tau - s|}\} \ \textit{(M)} \end{aligned}$$

and no (T). (M) can be rewritten

$$\alpha(s)(p_2(s) + \sqrt{|\tau - s|}) = \min_{a \in [-1,1]} \{a(p_2(s) + \sqrt{|\tau - s|})\}$$

By (ADJ) p_2 is linear. Since the square root grows faster at the vicinity of τ ($p_2(t)$ must also be zero at τ, otherwise it is clearly not a switching point for control coverned by (M)), then $p_2(s) + \sqrt{|\tau - s|}$ is positive in $U \setminus \{\tau\}$

for some neighborhood U, and this again contradicts the switching property of α.

We observe that the usual Pontryagin (without recalling the extra 0 multiplier) produces a false control.

4. Infinite time horizon

Use the standing assumptions. Consider problem of minimizing

$$u(x) := \inf_{\alpha} C_x(\alpha) = \inf_{\alpha} \int_0^\infty r(x(s), \alpha(s)) e^{-\lambda s}\, ds$$

$\lambda > 0$, with otherwise the same setup as before. Then u bounded, and Lip or Hölder continuous depending on the ratio of λ/L, where L is Lip constant of dynamics f. See [BCD97] Prop III.2.1.

Theorem 4.1 (DPP)**.** *Let u be a value, $(x,t) \in \mathbb{R}^n \times (0,\infty)$, and standing assumptions hold. Then*

$$u(x) = \inf_{\alpha \in \mathcal{A}} \{ \int_0^t r(x(s), \alpha(s)) e^{-\lambda s}\, ds + e^{-\lambda t} u(x(t)) \}$$

Proof. Aim: Define a control

$$\alpha(s) = \begin{cases} \alpha_1(s), & 0 \le s \le t \\ \alpha_2(s), & t < s. \end{cases}$$

Given α_1 and related trajectory $x_1(s)$ with $x_1(t) = x$, choose α_2 (and related trajectory with $x_2(t) = x_1(t)$) so that

$$\int_t^\infty r(x_2(s), \alpha_2(s)) e^{-\lambda(s-t)}\, ds - \eta \le u(x_1(t))$$

which implies

$$\begin{aligned} u(x) &\overset{\text{use } \alpha}{\le} \int_0^t r(x_1(s), \alpha_1(s)) e^{-\lambda s}\, ds + e^{-\lambda t} \int_t^\infty r(x_2(s), \alpha_2(s)) e^{-\lambda(s-t)}\, ds \\ &\le \int_0^t r(x_1(s), \alpha_1(s)) e^{-\lambda s}\, ds + e^{-\lambda t} u(x_1(t)) + \eta. \end{aligned}$$

Taking $\inf_{\alpha_1}$ implies

$$u(x) \le w(x),$$

where $w(x) := \inf_{\alpha\in\mathcal{A}}\{\int_0^t r(x(s),\alpha(s))e^{-\lambda s}\,ds + e^{-\lambda t}u(x(t))\}$. To prove a reverse, choose α_3, $x_3(0) = x$ such that

$$\begin{aligned}
u(x) &\geq \int_0^\infty r(x_3(s),\alpha_3(s))e^{-\lambda s}\,ds - \eta \\
&\geq \int_0^t r(x_3(s),\alpha_3(s))e^{-\lambda s}\,ds + e^{-\lambda t}\int_t^\infty r(x_3(s),\alpha_3(s))e^{-\lambda(s-t)}\,ds - \eta \\
&\geq \int_0^t r(x_3(s),\alpha_3(s))e^{-\lambda s}\,ds + e^{-\lambda t}u(x_3(s)) - \eta. \\
&\geq w(x) - \eta.
\end{aligned}$$

□

We guess by formal derivation that the related HJB PDE is

$$\begin{aligned}
&\inf_{\alpha\in\mathcal{A}}\{\frac{1}{t}\int_0^t r(x(s),\alpha(s))e^{-\lambda s}\,ds + \frac{1}{t}(e^{-\lambda t}u(x(t)) - u(x))\} \\
&\quad \to \inf_{a\in A}\{r(x,a) - \lambda e^0 u(x) + e^0 Du(x)\cdot x'(0)\} \\
&\quad = \inf_{a\in A}\{r(x,a) + Du(x)\cdot f(x,a)\} - \lambda u(x) \\
&\quad = H(x, Du(x)) - \lambda u(x) = 0
\end{aligned}$$

The rigorous proof that value function is a viscosity solution runs along the same lines as before.

5. Problems with exit times

Under the standing assumptions. Consider a problem of minimizing

$$C_x(\alpha) = \begin{cases} \int_0^\tau r(x(s),\alpha(s))e^{-\lambda s}\,ds + e^{-\lambda\tau}g(x(\tau)), & \tau < \infty \\ \int_0^\infty r(x(s),\alpha(s))e^{-\lambda s}\,ds, & \tau = \infty, \end{cases}$$

where $\tau := \tau(\alpha, x)$ is the first hitting time a given closed target $\mathcal{T}$ with compact boundary $\partial\mathcal{T}$, and we assume $r \geq 1$ and $\lambda \geq 0$, and if $\lambda = 0$, then we assume $\inf_\alpha \tau < \infty$ for every point, in addition to the usual standing assumptions.

Theorem 5.1 (DPP)**.** *Let u be a value, $x \in \mathbb{R}^n \setminus \mathcal{T}$, and the above assumptions hold. Then*

$$\begin{aligned}
u(x) = \inf_{\alpha\in\mathcal{A}}\{\int_0^{t\wedge\tau} r(x(s),\alpha(s))e^{-\lambda s}\,ds + \chi_{\{\tau\leq t\}}e^{-\lambda\tau}g(x(\tau)) \\
+ \chi_{\{\tau>t\}}e^{-\lambda t}g(x(t))\}.
\end{aligned}$$

The proof uses the usual techniques.

Again formally, for $x \in \mathbb{R}^n \setminus \mathcal{T}$,

$$\begin{aligned}
0 &= \inf_{\alpha\in\mathcal{A}}\Big\{\frac{1}{t\wedge\tau}\int_0^{t\wedge\tau} r(x(s),\alpha(s))e^{-\lambda s}\,ds + \frac{1}{t\wedge\tau}\chi_{\{\tau\le t\}}(e^{-\lambda\tau}g(x(\tau)) - u(x)) \\
&\qquad + \frac{1}{t\wedge\tau}\chi_{\{\tau>t\}}(e^{-\lambda t}u(x(t)) - u(x))\Big\} \\
&\to \inf_{a\in A}\{r(x,a) - \lambda u(x) + Du(x)\cdot f(x,a)\} \\
&= \inf_{a\in A}\{r(x,a) + Du(x)\cdot f(x,a)\} - \lambda u(x) \\
&=: H(x, Du(x,t)) - \lambda u(x)
\end{aligned}$$

since eventually $t < \tau$.

If the value u is continuous up to the $\partial\mathcal{T}$ (we do not pursue this question further), in addition to the above assumptions, then u is a viscosity solution to the Dirichlet problem

$$\begin{cases} H(x, Du(x,t)) - \lambda u(x) = 0, & \text{in } \mathbb{R}^n \setminus \mathcal{T}, \\ u(x) = g(x), & \text{on } \partial\mathcal{T}. \end{cases}$$

Example 5.2 (Warning)**.** *Consider a ball $B(0,1)$, the dynamics*

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

where $\alpha : [0,\tau] \to \overline{B}(0,1)$. This violates $r \geq 1$. Suppose we pay according to the path length i.e. minimize

$$P_x(\alpha) = \int_0^\tau |\alpha(s)|\,ds + 0.$$

Never moving gives a value 0, even if we take some bdr values $g > 0$. The corresponding HJB is

$$\begin{aligned}
&H(x, Du) \\
&\quad = \inf_{a\in\overline{B}(0,1)}\{f(x,a)\cdot Du(x) + r(x,a)\} \\
&\quad = \inf_{a\in\overline{B}}\{a\cdot Du(x) + |a|\} \\
&\quad = \min\{0, -|Du| + 1\} = 0,
\end{aligned}$$

but this has no unique solution.

The DPP considerations also work for the constrained endpoint problems.

Example 5.3 (Rocket car)**.** *Recall the formulation*

$$x(t) = \begin{pmatrix} q(t) \\ v(t) \end{pmatrix}$$

as

$$x'(t) = \begin{pmatrix} q'(t) \\ v'(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} x(t) + \begin{pmatrix} 0 \\ \alpha(t) \end{pmatrix}.$$

Denote

$$\begin{pmatrix} q \\ v \end{pmatrix} =: \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} =: x, \quad f(x,a) = \begin{pmatrix} x_2 \\ a \end{pmatrix},$$

and recall that aim was take rocket car to $q = 0, v = 0$ in minimum time. Denote the minimum time starting at x by $u(x)$. Dynamic programming principle:

$$u(x) = \inf_{\alpha \in \mathcal{A}} \{ \int_0^{\tau \wedge t} 1 \, ds + \chi_{\{\tau > t\}} u(x(t)) \}.$$

The Hamilton-Jacobi-Bellman equation is now

$$\begin{aligned} 0 &= \min_{a \in [-1,1]} \{ r(x,a) + Du(x) \cdot f(x,a) \} \\ &= \min_{a \in [-1,1]} \{ 1 + Du(x) \cdot f(x,a) \} \\ &= \min_{a \in [-1,1]} \{ 1 + Du(x) \cdot \begin{pmatrix} x_2 \\ a \end{pmatrix} \} \\ &= \min_{a \in [-1,1]} \{ 1 + x_2 u_{x_1}(x) + a u_{x_2}(x) \}. \end{aligned}$$

6. Existence of optimal controls

We denote

$$L^\infty = L^\infty(0,t;\mathbb{R}^m) = \{ \alpha : [0,t] \to \mathbb{R}^m \; : \; \operatorname*{ess\,sup}_{0 \le s \le t} |\alpha(s)| < \infty \},$$

$$||\alpha||_{L^\infty} = \operatorname*{ess\,sup}_{0 \le s \le t} |\alpha(s)| \, .$$

Definition 6.1. *Let $\alpha_i, \alpha \in L^\infty$. We say that α_i converges to α in weak* if*

$$\int_0^t \alpha_i(s) \cdot v(s) \, ds \to \int_0^t \alpha(s) \cdot v(s) \, ds$$

as $i \to \infty$, for any $v : [0,t] \to \mathbb{R}^m$ with $\int_0^t |v(s)| \; ds < \infty$.

Theorem 6.2 (Alaoglu's theorem)**.** *Let $||\alpha_i||_{L^\infty} \le M < \infty$. Then passing to a subsequence if necessary, there is $\alpha \in L^\infty$ such that*

$$\int_0^t \alpha_i(s) \cdot v(s) \, ds \to \int_0^t \alpha(s) \cdot v(s) \, ds$$

as $i \to \infty$, for any $v : (0,t) \to \mathbb{R}^n$ with $\int_0^t |v(s)| \; ds < \infty$.

Definition 6.3. *A point $z \in K$ is extreme, if there are no $x, y \in K$ and $\lambda \in (0,1)$ such that*

$$z = \lambda x + (1 - \lambda) y \in K.$$

Theorem 6.4 (Krein-Milman)**.** *Let $\emptyset \neq K \subset L^\infty$ be convex and let $a_i \subset K$. If, passing to a subsequence if necessary, there is $\alpha \in L^\infty$ such that*

$$\int_0^t \alpha_i(s) \cdot v(s)\, ds \to \int_0^t \alpha(s) \cdot v(s)\, ds$$

as $i \to \infty$, for any $v : (0,t) \to \mathbb{R}^n$ with $\int_0^t |v(s)|\, ds < \infty$, then K has at least one extreme point.

We say that a control is of bang-bang type if (recall the rocket car example) if for each time $s \in [0,t]$ and $i = 1, 2, \dots m$, it holds that

$$|\alpha_i(s)| = 1.$$

Next we consider existence of the optimal bang-bang control in an example.

Example 6.5 (Rocket car)**.** *Recall that the dynamics in the rocket car example is of the form*

$$\begin{cases} x'(s) = Mx(s) + N\alpha(s), \\ x(0) = x_0. \end{cases}$$

Next we sketch the proof that there exists an optimal bang-bang control. Let

$$\tilde{\tau} = \inf\{t : \textit{ we can steer from } x_0 \textit{ to } 0 \textit{ in time } t \}.$$

We want to show that there is $\tilde{\alpha}$ steering from x_0 to 0 in time $\tilde{\tau}$. After earlier considerations, we take for granted that always $\tilde{\tau} < \infty$, and by definition of inf

$$t_i \to \tilde{\tau}+$$

as $i \to \infty$.

We can define $\alpha_i = 0$ after t_i so that all the controls are defined up to time t_1. By Banach-Alaoglu's Theorem, Theorem 6.2, passing to a subsequence if necessary there is $\tilde{\alpha}$ so that α_i converges in weak*.

Since by denoting $X(s) = e^{tM}$ it holds that (ex)

$$\begin{aligned} 0 &= X(t_1)x_0 + X(t_1)\int_0^{t_1} X(-s)N\alpha_i(s)\, ds \\ &\overset{\text{by weak}^* \text{ conv}}{\to} X(\tilde{\tau})x_0 + X(\tilde{\tau})\int_0^{\tilde{\tau}} X(-s)N\tilde{\alpha}(s)\, ds \end{aligned}$$

since $\int_0^{t_1} |X(-s)N|\, ds < \infty$, it follows that $\tilde{\alpha}$ is optimal. (This formula tells if we can reach 0 from x_0 in given time and control, and is based on the theory of linear ODEs, see Evans' lecture note.)

After verifying that set of controls steering to 0 in time $\tilde{\tau}$ satisfies the condition of Krein-Milman (ex, based on the above integral formula), we *can show that there is optimal control of bang-bang type: Let α^* be extreme point of controls steering to 0 in time $\tilde{\tau}$ (actually essentially $\alpha^* = \tilde{\alpha}$ in this particular example). Now we show that such extreme point is bang-bang. Suppose that there is $F \subset (0, \tilde{\tau}), |F| > 0$ such that*

$$\alpha^* < 1 - \varepsilon.$$

Now, choose $\beta \not\equiv 0$, $|\beta| \leq 1$, $\beta = 0$ outside F, and

$$\int_F X^{-1}(s) N \beta(s)\, ds = 0.$$

It holds that

$$|\alpha^* + \varepsilon\beta| \leq 1.$$

Thus since

$$\begin{aligned} &X(\tilde{\tau})x_0 + X(\tilde{\tau}) \int_F X^{-1}(s) N(\alpha^*(s) + \varepsilon\beta(s))\, ds \\ &= X(\tilde{\tau})x_0 + X(\tilde{\tau}) \Big(\int_F X^{-1}(s) N\alpha^*(s)\, ds + \varepsilon \underbrace{\int_F X^{-1}(s) N\beta(s)\, ds}_{=0} \Big) = 0 \end{aligned}$$

also $\alpha^* + \varepsilon\beta$ is admissible and steers to 0 in time $\tilde{\tau}$. So is $\alpha^* - \varepsilon\beta$, and $\frac{1}{2}(\alpha^* + \varepsilon\beta) + \frac{1}{2}(\alpha^* - \varepsilon\beta) = \alpha^*$, which is a contradiction, since α^* was extreme. Thus α^* must be of bang-bang type.

7. Stochastic control theory

This section gives some ideas on the stochastic control theory. It is strictly formal, for the backround and rigor, the stochastics courses are recommended.

So far our dynamics has been given by ODE,

$$\begin{cases} x'(t) = f(x(t), \alpha(t)), & t > 0 \\ x(0) = x_0. \end{cases}$$

In many financial etc. phenomenons, there is however noise and stochastic differential equations (SDE) are used as a model dynamics, for example,

$$\begin{cases} X'(t) = f(X(t), A(t)) + \text{noise} & t > 0 \\ X(0) = x_0. \end{cases}$$

and A is a control. Here X is an example of a **stochastic process**.

We skip the rigorous assumptions: roughly speaking, we need regularity assumptions on f and measurability assumptions on A. Heuristically

$$A : [0, T] \to K$$

is a mapping such that $A(s)$ depends on the history/observations up to time s.

What incidence of history occurred can be indexed by sample point $\omega \in \Omega$. Further, we need to assign sample space, events and their probabilities. This is modeled by **probability space**

$$(\Omega, \mathcal{F}, \mathbb{P}),$$

where $\mathcal{F}$ (a σ-algebra) is space of events and $\mathbb{P}$ a probability measure assigning probability $\mathbb{P}(A)$ on each $A \in \mathcal{F}$.

Random variables are usually denoted by capital letters and they are defined to be measurable functions with respect to the $\mathcal{F}$. This allows us define the **expectation** as a usual integral

$$\mathbb{E}[X] := \int_\Omega X \, d\mathbb{P}.$$

A key concept is so called **stochastic integral** with respect to Brownian motion W denoted

$$\int_0^T G \, dW.$$

This integral is being build on by approximating by step processes and sums like

$$\sum_k G_k(W(s_{k+1}) - W(s_k)).$$

A typical notation in SDE:s now reads as

$$\begin{cases} dX(t) = f(t)\, dt + \sigma(t)\, dW, \\ X(0) = x_0 \end{cases}$$

where $f : [0, T] \to \mathbb{R}^n$ and $\sigma : [0, T] \to \mathbb{R}$ (we take σ to be deterministic scalar valued for simplicity). This means

$$X(t) = x_0 + \int_0^t f(s)\, ds + \int_0^t \sigma(s)\, dW(s).$$

Heuristically, **Ito rule** can looked through Taylor's theorem. A difference is that the Brownian motion is very 'curly' or irregular and thus also higher order effects add up. In 1D, assuming u is smooth,

$$\begin{aligned} du(X(t)) &= u'(X(t))\, dX(t) + \frac{1}{2} u''(X(t))(\, dX)^2 + \dots \\ &= u'(X(t))(f(t)\, dt + \sigma(t)\, dW) + \frac{1}{2} u''(X(t))(f(t)\, dt + \sigma(t)\, dW)^2 + \dots \\ &= u'(X(t))(f(t)\, dt + \sigma(t)\, dW) \\ &\quad + \frac{1}{2} u''(X(t))(f^2(t)(\, dt)^2 + 2f(t)\sigma(t)\, dt\, dW + \sigma^2(t)(\, dW(t))^2 + \dots \end{aligned}$$

Using $dW^2 = dt$ and dropping the higher order terms, we obtain

$$du(X(t)) = \Big(u'(X(t))f(t) + \frac{1}{2} u''(X(t))\sigma^2(t)\Big)\, dt + u'(X(t))\sigma(t)\, dW.$$

In the higher order case, where dX, $dW = (dW_1, \dots, dW_n)$ etc are vectors. We use the rule

$$dW_i\, dW_j = \begin{cases} dt, & i = j \\ 0, & \text{otherwise.} \end{cases}$$

and get, also including time dependence for u,

$$\begin{aligned}
&du(X(t), t) \\
&= u_t(X(t), t)\, dt + \sum_i u_{x_i}(X(t))\, dX_i(t) + \frac{1}{2} \sum_{ij} u_{x_i x_j}(X(t))\, dX_i\, dX_j \\
&= u_t(X(t), t) + \sum_i u_{x_i}(X(t))(f_i(t)\, dt + \sigma(t)\, dW_i(t)) \\
&\quad + \frac{1}{2} \sum_{ij} u_{x_i x_j}(X(t))(f_i(t)\, dt + \sigma\, dW_i)(f_j(t)\, dt + \sigma(t)\, dW_j(t)) \\
&= u_t(X(t), t)\, dt + \sum_i \left(\left(u_{x_i}(X(t), t) f_i(t) + u_{x_i x_i}(X(t), t) \frac{\sigma(t)^2}{2} \right) dt + \sigma(t) u_{x_i}(X(t), t)\, dW_i(t) \right) \\
&= \left(u_t(X(t), t) + Du(X(t), t) \cdot f(t) + \Delta u(X(t), t) \frac{\sigma(t)^2}{2} \right) dt + \sigma(t) Du(X(t), t) \cdot dW(t).
\end{aligned}$$

7.1. **Dynamic programming and connection to PDE.** Consider the controlled SDE

$$\begin{cases} dX(s) = f(X(s), A(s))\, ds + \sigma\, dW(s), & t \le s \le T \\ X(t) = x. \end{cases}$$

We define the expected payoff

$$P_{x,t}(A) := \mathbb{E}\left[\int_t^T r(X(s), A(s))\, ds + g(X(T)) \right]$$

and the value

$$\begin{aligned}
u(x, t) &= \sup_{A \in \mathcal{A}} P_{x,t}(\mathcal{A}) \\
&= \mathbb{E}\left[\int_t^T r(X(s), A(s))\, ds + g(X(T)). \right]
\end{aligned}$$

Let A be

$$A(s) = \begin{cases} \text{any control}, & t \le s \le t + h \\ \text{optimal}, & t + h < s \le T. \end{cases}$$

Since the value u is the largest expected payoff, we have

$$u(x, t) \ge \mathbb{E}\left[\int_t^{t+h} r(X(s), A(s))\, ds + u(X(t + h), t + h) \right].$$

Thus

$$0 \geq \mathbb{E}\left[\int_t^{t+h} r(X(s), A(s))\, ds\right] + \mathbb{E}\Big[u(X(t+h), t+h) - u(x,t)\Big] \tag{7.16}$$

Recalling Ito formula (replace A there by f) and writing with the integral notation, we have

$$\begin{aligned} &u(X(t+h), t+h) - u(X(t), t) \\ &\quad = \int_t^{t+h} \big(u_t + Du \cdot f + \Delta u \frac{\sigma^2}{2}\big)\, ds + \sigma \int_t^{t+h} Du \cdot dW \end{aligned}$$

so that

$$\begin{aligned} &\mathbb{E}\Big[u(X(t+h), t+h) - u(x,t)\Big] \\ &\quad = \mathbb{E}\left[\int_t^{t+h} \big(u_t + Du \cdot f + \Delta u \frac{\sigma^2}{2}\big)\, ds + \sigma \int_t^{t+h} Du \cdot dW\right] \\ &\quad = \mathbb{E}\left[\int_t^{t+h} \big(u_t + Du \cdot f + \Delta u \frac{\sigma^2}{2}\big)\, ds\right] + \sigma \underbrace{\mathbb{E}\left[\int_t^{t+h} Du \cdot dW\right]}_{=0}. \end{aligned}$$

Combining this with (7.16), we have

$$0 \geq \mathbb{E}\left[\int_t^{t+h} \big(r + u_t + Du \cdot f + \Delta u \frac{\sigma^2}{2}\big)\, ds\right].$$

Assume that $X(s) \to x, A(s) \to a$ as $s \to 0$, f, r and u are regular, divide by h and pass to a limit to get

$$0 \geq r(x,a) + u_t(x,t) + Du(x,t) \cdot f(x,a) + \Delta u(x,t) \frac{\sigma^2}{2}. \tag{7.17}$$

Since this holds for any control up to time $t+h$ and we expect equality if taking sup over controls

$$\begin{aligned} 0 = \sup_{A \in \mathcal{A}} \mathbb{E}\Bigg[\int_t^{t+h} \Bigg(& r(X(s), A(s)) + u_t(X(s), s) + Du(X(s), s) \cdot f(X(s), A(s)) \\ &+ \Delta u(X(s), s) \frac{\sigma^2}{2}\Bigg)\, ds\Bigg] \end{aligned}$$

Thus there is a control s.t.

$$\begin{aligned} 0 \leq \mathbb{E}\Bigg[\int_t^{t+h} \Bigg(& r(X(s), A(s)) + u_t(X(s), s) + Du(X(s), s) \cdot f(X(s), A(s)) \\ &+ \Delta u(X(s), s) \frac{\sigma^2}{2}\Bigg)\, ds\Bigg] - \eta. \end{aligned}$$

Assuming that $X(s) \to x, A(s) \to a$ as $s \to 0$, f, r and u are regular, divide by h and pass to a limit we get

$$0 \leq r(x,a) + u_t(x,t) + Du(x,t) \cdot f(x,a) + \Delta u(x,t) - \eta. \tag{7.18}$$

Since $\eta > 0$ above is arbitrary and recalling (7.17), we justified heuristically that the value satisfies a PDE

$$u_t(x,t) + \sup_{a\in K}\{r(x,a) + Du(x,t)\cdot f(x,a)\} + \Delta u(x,t) = 0.$$

Remark 7.1. *Now the highest order operator has no control dependence since σ was just a constant. This would change by choosing σ with control dependence.*

Example 7.2 (Merton's optimal portfolio selection problem)**.** *Next we look at Application 7.6 in Evans' lecture note (done at lectures). As shown there the solution to the corresponding HJB is of the form $u(x,t) = g(t)x^\gamma$, where x reflects the wealth, and γ is the power in the selected utility function for consumption. Also g is explicitly given in Evans' lecture note. The optimal portfolio selection is*

$$\begin{aligned}\tilde{\alpha}_1(x,t) &= \frac{-(R-r)u_x}{\sigma^2 x u_{xx}}\\ &= \frac{-(R-r)\gamma x^{\gamma-1}}{\sigma^2 x\gamma(\gamma-1)x^{\gamma-2}}\\ &= \frac{-(R-r)\gamma x^{\gamma-1}}{\sigma^2 x\gamma(\gamma-1)x^{\gamma-2}}\\ &= \frac{-(R-r)}{\sigma^2(\gamma-1)}.\end{aligned}$$

References

[ABIL13] Y. Achdou, G. Barles, H. Ishii, and G. L. Litvinov. *Hamilton-Jacobi Equations: Approximations, Numerical Analysis and Applications: Cetraro, Italy 2011, Editors: Paola Loreti, Nicoletta Anna Tchou*, volume 2074. Springer, 2013.

[Bar94] Guy Barles. *Solutions de viscosité des équations de Hamilton-Jacobi*. 1994.

[BCD97] M. Bardi and I. Capuzzo-Dolcetta. Optimal control and viscosity solutions of hamilton-jacobi-bellman equations. 1997.

[Eva05] L. C. Evans. An introduction to mathematical optimal control theory. *Lecture Notes, University of California, Department of Mathematics, Berkeley*, 2005.

[Eva10] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.

[FS06] W. H. Fleming and H. M. Soner. *Controlled Markov processes and viscosity solutions*, volume 25 of *Stochastic Modelling and Applied Probability*. Springer, New York, second edition, 2006.

[Kry09] N. V. Krylov. *Controlled diffusion processes*, volume 14 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2009. Translated from the 1977 Russian original by A. B. Aries, Reprint of the 1980 edition.

[Lio82] P.-L. Lions. *Generalized solutions of Hamilton-Jacobi equations*, volume 69. Pitman Publishing, 1982.

[Øks03] B. Øksendal. *Stochastic differential equations*. Universitext. Springer-Verlag, Berlin, sixth edition, 2003. An introduction with applications.

[Tou12] N. Touzi. *Optimal stochastic control, stochastic target problems, and backward SDE*, volume 29. Springer Science & Business Media, 2012.

Department of Mathematics and Statistics, University of Jyväskylä, PO Box 35, FI-40014 Jyväskylä, Finland

E-mail address: Please email typos/errors to:mikko.j.parviainen@jyu.fi