LECTURE 5

5. The Hamilton-Jacobi-Bellman equation

At the same time that the maximum principle was developed in Soviet, a fundamentally different method for solving the same type of problems was developed in the west. This method, which seems more intuitive, is based on what is called *the principle of optimality* of dynamic programming.

5.1. **Dynamic Programming.**

To illustrate dynamic programming let us consider the following example:

Example Given a function $g : [0, a] \to \mathbb{R}_+$, find the partition of the interval $[0, a]$ into N different subintervals $[0, x_1], [x_1, x_2], \dots, [x_{N-1}, a]$ that maximizes

$$\frac{1}{2}\sum_{k=1}^{N} g\left(\frac{x_{k-1}+x_k}{2}\right)(x_k - x_{k-1}),$$

with $x_0 = 0$ and $x_N = a$.

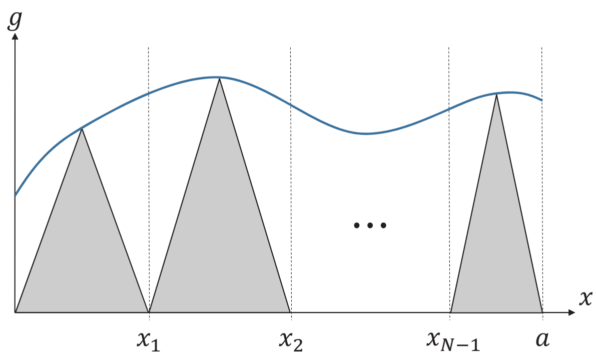

Figure 1. The partition in the Example.

To solve this problem we let $V_{k-1}(x)$ denote the optimal value for a partition of $[0, x]$ into $k-1$ intervals. Adding one more interval at the end, $[x-u, x]$, we get

$$V_k(x) = \sup_{0\le u\le x}\left\{\frac{1}{2}g(x-u/2)u + V_{k-1}(x-u)\right\},$$

with

$$V_1(x) = \frac{1}{2}g(x/2)x.$$

As an example let $g(x) = x$ and let $a = 1$, then

$$V_1(x) = \frac{1}{4}x^2$$

and

$$\begin{aligned} V_2(x) &= \sup_{0\le u\le x}\left\{\frac{1}{2}g(x-u/2)u + V_1(x-u)\right\} \\ &= \sup_{0\le u\le x}\left\{\frac{1}{2}(x-u/2)u + \frac{1}{4}(x-u)^2\right\} \\ &= \sup_{0\le u\le x}\left\{\frac{1}{4}\left(2xu - u^2 + x^2 - 2xu + u^2\right)\right\} \\ &= \frac{1}{4}x^2. \end{aligned}$$

Hence, by induction $V_N(x) = \frac{1}{4}x^2$, giving us $V_N(1) = 1/4$ independent of the partition. □

We now consider the target set $S = \{t_1\} \times \mathbb{R}^n$ and consider the family of problems

$$J(t,x,u) = \int_t^{t_1} L(s,x(s),u(s))ds + K(x(t_1)),$$

where $x(t) = x$, and define the *value function*

$$V(t,x) := \inf_{u_{[t,t_1]}} J(t,x,u),$$

where $u_{[t,t_1]}$ is the set of controls restricted to $[t,t_1]$.

$V(t,x)$ is thus the optimal "to go" cost from t to t_1 when $x(t) = x$. The end-time condition gives that $V(t_1,x) = K(x)$.

Bellman's approach to solve optimal control problems was to derive a PDE for the value function through the principle of optimality.

5.2. **Principle of optimality.** The principle of optimality says that an optimal trajectory must be optimal everywhere. If we partition the interval $[t,t_1]$ into two intervals $[t,t+\Delta t]$ and $[t+\Delta t,t_1]$. Then the control must be optimal on $[t+\Delta t,t_1]$ with $K(x(t_1))$ as terminal cost, and optimal on $[t,t+\Delta t]$ with $V(t+\Delta t, x(t+\Delta t))$ as terminal cost. Hence,

$$V(t,x) = \inf_{u_{[t,t+\Delta t]}} \int_t^{t+\Delta t} L(s,x(s),u(s))ds + V(t+\Delta t, x(t+\Delta t)),$$

for every $(t,x) \in [t_0,t_1]\times\mathbb{R}^n$ and $\Delta t \in (0, t_1 - t]$.

5.3. **The HJB equation.** We have that $x(t+\Delta t) \approx x + f(t,x,u(t))\Delta t$ which gives

$$V(t+\Delta t, x(t+\Delta t)) \approx V(t,x) + V_t(t,x)\Delta t + \langle V_x, f(t,x,u(t))\rangle\Delta t.$$

We also have that

$$\int_t^{t+\Delta t} L(s,x(s),u(s))ds \approx L(t,x,u(t))\Delta t.$$

Putting this together we get that

$$V(t,x) \approx \inf_{u_{[t,t+\Delta t]}} \left\{L(t,x,u(t))\Delta t + V(t,x) + V_t(t,x)\Delta t + \langle V_x(t,x), f(t,x,u(t))\rangle\Delta t\right\}.$$

We observe here that the $V(t,x)$ cancel out and we get

$$0 \approx \inf_{u_{[t,t+\Delta t]}} \left\{L(t,x,u(t))\Delta t + V_t(t,x)\Delta t + \langle V_x(t,x), f(t,x,u(t))\rangle\Delta t\right\}.$$

Dividing by Δt and letting $\Delta t \to 0$ we get

$$0 = \inf_{u\in U}\left\{L(t,x,u) + V_t(t,x) + \langle V_x(t,x), f(t,x,u)\rangle\right\}.$$

This equation is called the Hamilton-Jacobi-Bellman (HJB) equation and is often written

$$\boxed{-V_t(t,x) = \inf_{u\in U}\left\{L(t,x,u) + \langle V_x(t,x), f(t,x,u)\rangle\right\}.}$$

The HJB equation is a PDE for the value function, as opposed to the maximum principle which gave us an ODE, with boundary condition given by $V(t_1,x) = K(x(t_1))$.

The HJB equation can also be written

$$V_t(t,x) = \sup_{u\in U}\left\{\langle -V_x(t,x), f(t,x,u)\rangle - L(t,x,u)\right\},$$

which can be compared to

$$H(t,x,u,p) = \langle p, f(t,x,u)\rangle - L(t,x,u).$$

This brings us to the Hamiltonian form of the HJB equation:

$$V_t(t,x) = \sup_{u\in U} H(t,x,u,-V_x(t,x)).$$

So far we have only shown necessity, if an optimal control exists then the corresponding value function will fulfill the HJB equation. However, sufficiency can also be shown:

Suppose that a $\mathcal{C}^1$ function $\hat{V} : [t_0,t_1]\times R^n \to \mathbb{R}$ satisfies the HJB equation:

$$-\hat{V}_t(t,x) = \inf_{u\in U}\left\{L(t,x,u) + \langle \hat{V}_x(t,x), f(t,x,u)\rangle\right\},$$

for all $t\in[t_0,t_1)$ and all $x\in\mathbb{R}^n$ and the boundary condition

$$\hat{V}(t_1,x) = K(x(t_1)).$$

Suppose that a control $\hat{u} : [t_0,t_1]\to U$ and the corresponding trajectory $\hat{x} : [t_0,t_1]\to\mathbb{R}^n$, with the given initial condition $\hat{x}(t_0) = x_0$, satisfy everywhere the equation

$$L(t,\hat{x}(t),\hat{u}(t)) + \langle \hat{V}_x(t,\hat{x}(t)), f(t,\hat{x}(t),\hat{u}(t))\rangle = \inf_{u\in U}\left\{L(t,\hat{x}(t),u) + \langle \hat{V}_x(t,\hat{x}(t)), f(t,\hat{x}(t),u)\rangle\right\}.$$

Then $\hat{V}(t_0,x_0)$ is the optimal cost and $\hat{u}$ is the optimal control.

To prove this note that since

$$-\hat{V}_t(t,\hat{x}(t)) = L(t,\hat{x}(t),\hat{u}(t)) + \langle \hat{V}_x(t,\hat{x}(t)), f(t,\hat{x}(t),\hat{u}(t))\rangle$$

we have

$$0 = L(t,\hat{x}(t),\hat{u}(t)) + \frac{d}{dt}\hat{V}(t,\hat{x}(t)).$$

Hence,

$$0 = \int_{t_0}^{t_1} L(t,\hat{x}(t),\hat{u}(t))dt + \hat{V}(t_1,\hat{x}(t_1)) - \hat{V}(t_0,\hat{x}(t_0)),$$

or

$$\hat{V}(t_0,\hat{x}(t_0)) = \int_{t_0}^{t_1} L(t,\hat{x}(t),\hat{u}(t))dt + K(\hat{x}(t_1)) = J(t_0,x_0,\hat{u}).$$

For any other control u with corresponding trajectory x with the same initial conditions we have

$$-\hat{V}_t(t,x(t)) \le L(t,x(t),u(t)) + \langle \hat{V}_x(t,x(t)), f(t,x(t),u(t))\rangle$$

or

$$0 \le L(t,x(t),u(t)) + \frac{d}{dt}\hat{V}(t,x(t)).$$

Integrating this gives us

$$\hat{V}(t_0,\hat{x}(t_0)) \le \int_{t_0}^{t_1} L(t,x(t),u(t))dt + K(x(t_1)) = J(t_0,x_0,u).$$

Hence, $\hat{V}(t_0, \hat{x}(t_0))$ is the minimal cost and $\hat{u}$ is an optimal control as no other control gives a lower cost. □

Note that $\hat{u}$ gives a global minimum as opposed to the maximum principle which gave a necessary condition for local optimality.

5.4. **Viscosity solutions. Example** Consider the problem

$$\min_u x(t_1) \quad \text{subj. to} \quad \begin{cases} \dot{x}(t) = u(t)x(t), & u(t) \in [-1, 1], \\ x(0) = x_0, & x(t_1) \text{ free.} \end{cases}$$

By inspection we find that the optimal control for this problem is $u \equiv -1$ if $x_0 > 0$ and $u \equiv 1$ if $x_0 < 0$. The value function is

$$V(t, x) = \begin{cases} xe^{-(t_1 - t)} & \text{if } x > 0, \\ xe^{t_1 - t} & \text{if } x < 0, \\ 0 & \text{if } x = 0. \end{cases}$$

The HJB equation reads $-V_t = \inf_{u \in [-1,1]} \{V_x x u\} = -|xV_x|$ with boundary condition $V(t_1, x) = x$.

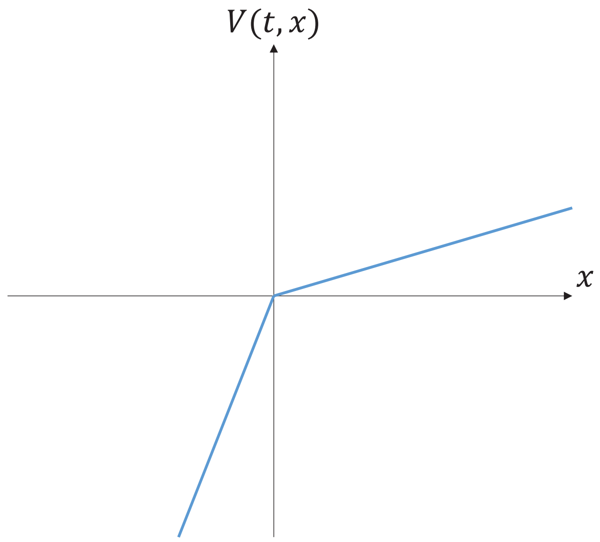

Figure 2. A value function that is not $\mathcal{C}^1$.

This example shows that there are optimal control problems which have value functions that are not $\mathcal{C}^1$. Hence, they cannot solve the HJB equation at all points in the classical sense. However, it can be shown that they still represent something called *viscosity solutions* to their HJB equations.

5.4.1. *Super- and sub-differentials.* Let $v : \mathbb{R}^n \to \mathbb{R}$ be a continuous function. A vector $\xi \in \mathbb{R}^n$ is called a *super-differential* to v in x if, for every y close to x, we have

$$v(y) \leq v(x) + \langle (y - x), \xi \rangle + o(|y - x|).$$

A *sub-differential* similarly fulfills

$$v(y) \geq v(x) + \langle (y - x), \xi \rangle - o(|y - x|).$$

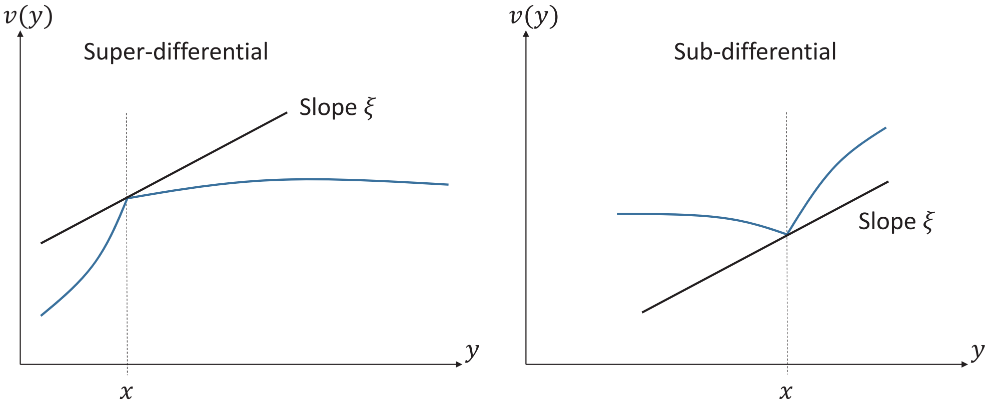

FIGURE 3. Super- and Sub-differentials.

The set of super-differentials at x is denoted $D^+v(x)$ and the set of sub-differentials at x is denoted $D^-v(x)$.

Example Let

$$v(x) = \begin{cases} 0 & \text{if } x < 0, \\ \sqrt{x} & \text{if } 0 \le x \le 1, \\ 1 & \text{if } x > 1. \end{cases}$$

Then at $x = 0$ we have $D^+v(0) = \emptyset$ and $D^-v(0) = [0, \infty)$. At $x = 1$ we have $D^+v(1) = [0, 1/2]$ and $D^-v(1) = \emptyset$.

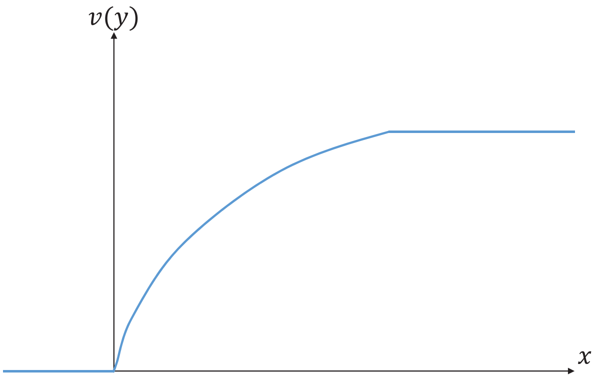

FIGURE 4. The function v in the example.

Claims:

(1) At all points x where v is $\mathcal{C}^1$, we have $D^+v(x) = D^-v(x) = \nabla v(x)$.
(2) If both $D^+v(x)$ and $D^-v(x)$ are non-empty, then $\nabla v(x)$ exists.

5.4.2. *Test functions.* $\xi \in D^+v(x)$ iff there is a $\mathcal{C}^1$ function $\varphi : \mathbb{R}^n \to \mathbb{R}$ such that $\nabla\varphi(x) = \xi$, $\varphi(x) = v(x)$ and $\varphi(y) \ge v(y)$, for all y sufficiently close to x, *i.e.* $\varphi - v$ has a local minimum at x.

Similarly, $\xi \in D^-v(x)$ iff there is a $\mathcal{C}^1$ function $\varphi : \mathbb{R}^n \to \mathbb{R}$ such that $\nabla\varphi(x) = \xi$, $\varphi(x) = v(x)$ and $\varphi(y) \le v(y)$, for all y sufficiently close to x, *i.e.* $\varphi - v$ has a local maximum at x.

The function φ is sometimes called a *test function*.

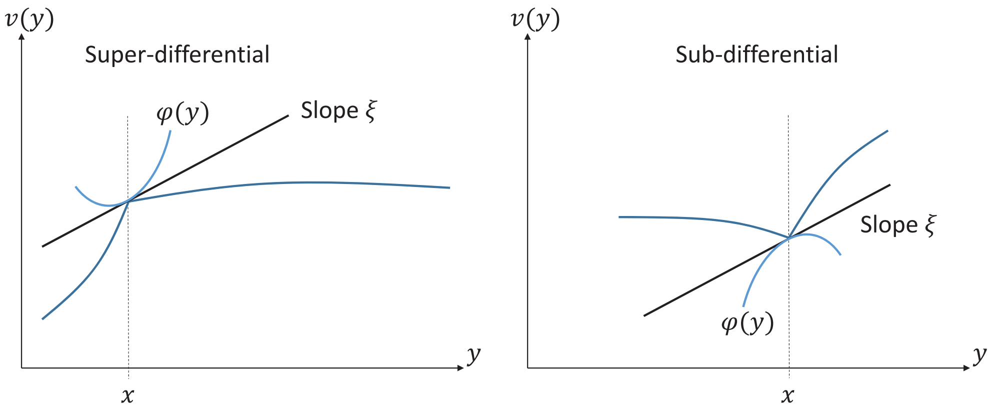

FIGURE 5. Test functions.

Using test functions we can prove the above claims:

(1) Let v be $\mathcal{C}^1$ at x. If $\varphi - v$ has a local minimum at x, then $\nabla(\varphi - v)(x) = 0$. Hence, $\nabla\varphi(x) = \nabla v(x)$ so that $D^+v(x) = \nabla v(x)$. Replacing minimum with maximum we get that $D^-v(x) = \nabla v(x)$.
(2) Assume that $\varphi_1, \varphi_2 \in \mathcal{C}^1$ are such that $\varphi_1(x) = \varphi_2(x) = v(x)$ and $\varphi_1(x) \le v(x) \le \varphi_2(x)$ close to x. Then $\varphi_1 - \varphi_2$ has a local minimum at x, hence $\nabla\varphi_1(x) = \nabla\varphi_2(x)$. By the "Sandwich theorem" v is then differentiable at x, and $\nabla v(x) = \nabla\varphi_1(x) = \nabla\varphi_2(x)$.

Using test functions it can also be shown that the sets $\{x : D^+v(x) \neq \emptyset\}$ and $\{x : D^-v(x) \neq \emptyset\}$ are both dense in the domain of v.

5.4.3. *Viscosity solutions of PDEs.* Consider a PDE of the form

$$F(x, v(x), \nabla v(x)) = 0,$$

where $F : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ is continuous.

A *viscosity subsolution* to the PDE is a continuous function $v : \mathbb{R}^n \to \mathbb{R}$, such that

$$F(x, v(x), \xi) \le 0 \quad \forall \xi \in D^+v(x),\ \forall x.$$

Which can be restated as $F(x, v(x), \nabla\varphi(x)) \le 0$ for all $\mathcal{C}^1$ test functions φ such that $\varphi - v$ has a local minimum at x. Similarly a *viscosity supersolution* to the PDE is a continuous function $v : \mathbb{R}^n \to \mathbb{R}$, such that

$$F(x, v(x), \xi) \ge 0 \quad \forall \xi \in D^-v(x),\ \forall x,$$

or equivalently $F(x, v(x), \nabla\varphi(x)) \ge 0$ for all $\mathcal{C}^1$ test functions φ such that $\varphi - v$ has a local maximum at x.

A *viscosity solution* of the PDE is both a viscosity sub- and a viscosity super-solution.

Example Let $F(x, v, \xi) = 1 - |\xi|$, which corresponds to the PDE $1 - |\nabla v(x)| = 0$. Both $v(x) = x$ and $v(x) = -x$ are classical solutions to this PDE. We claim that $v(x) = |x|$ is a viscosity solution.
We have that $D^+v(0) = \emptyset$, hence, the subsolution property holds trivially. For the supersolution part we note that $D^-v(0) = [-1, 1]$ and $F(0, 0, \xi) = 1 - |\xi| \ge 0$ for $\xi \in [-1, 1]$.

Note that the viscosity solution property is dependent on the sign on F. The term viscosity solution comes from the fact that the solution can be obtained as the limit $F(x, v_\varepsilon(x), \nabla v_\varepsilon(x)) = \varepsilon\Delta v_\varepsilon(x)$, when $\varepsilon \searrow 0$.

5.4.4. *Viscosity solutions of the HJB equation.*

Theorem 5.1. *The value function V is a unique viscosity solution to the HJB equation, with boundary conditions $V(t_1, x) = K(x)$, for all $x \in \mathbb{R}^n$.*

We prove the subsolution part and leave the proof of the supersolution property as a home assignment. We thus want to show that for every $\mathcal{C}^1$ test function φ such that $\varphi - V$ has a local minimum in (t_0, x_0),

$$-\varphi_t(t_0, x_0) - \inf_{u \in U} \Big\{ L(t_0, x_0, u) + \langle \varphi_x(t_0, x_0), f(t_0, x_0, u) \rangle \Big\} \leq 0.$$

Assume the opposite, *i.e.* that there is a $\mathcal{C}^1$ function φ and a control value u_0 such that

$$\varphi(t_0, x_0) = V(t_0, x_0), \quad \text{and} \quad \varphi(t, x) \geq V(t, x), \text{ near } (t_0, x_0),$$

and

$$-\varphi_t(t_0, x_0) - L(t_0, x_0, u_0) - \langle \varphi_x(t_0, x_0), f(t_0, x_0, u_0) \rangle > 0. \tag{5.1}$$

Now assume that we apply the control $u \equiv u_0$ on the short interval $[t_0, t_0 + \Delta t]$ and get the trajectory x. We have

$$\begin{aligned} V(t_0 + \Delta t, x(t_0 + \Delta t)) - V(t_0, x_0) &\leq \varphi(t_0 + \Delta t, x(t_0 + \Delta t)) - \varphi(t_0, x_0) \\ &= \int_{t_0}^{t_0 + \Delta t} \frac{d}{dt} \varphi(t, x(t)) dt \\ &= \int_{t_0}^{t_0 + \Delta t} \left(\varphi_t(t, x(t)) + \langle \varphi_x(t, x(t)), f(t, x(t), u_0) \rangle \right) dt \\ &< - \int_{t_0}^{t_0 + \Delta t} L(t, x(t), u_0) dt \end{aligned}$$

where the last step follows from (5.1) and continuity. This means that

$$V(t_0, x_0) - V(t_0 + \Delta t, x(t_0 + \Delta t)) > \int_{t_0}^{t_0 + \Delta t} L(t, x(t), u_0) dt.$$

Hence, the control $u \equiv u_0$ would outperform the optimal control on the interval $[t_0, t_0 + \Delta t]$, contradicting the principle of optimality. □