Dynamic programming approach for optimal control problems

Zhiping Rao

August 9-13, 2021

In this short course, we will first introduce the definition of the value function for the optimal control problems and the associated dynamic programming principle. Then we will derive the Hamilton-Jacobi-Bellman equation satisfied by the value function and discuss the theory of viscosity solution for this type of equation, including the existence, uniqueness and stability of the solution. Finally, we will give some numerical methods to compute the value function, including the finite difference schemes for Hamilton-Jacobi-Bellman equations and the Semi-Lagrangian scheme based on the dynamic programming principle.

1 Lecture 1

1.1 Optimal control problem

Given $T > 0$, for each $(t, x) \in [0, T] \times \mathbb{R}^n$, consider the following dynamical system:

$$\begin{cases} \dot{y}(s) = f(y(s), u(s)) \text{ a.e. } s \in (t, T), \\ y(t) = x, \end{cases} \tag{1.1}$$

where $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ and $u(\cdot) \in L^\infty(t, T; U)$ where U is a compact subset of $\mathbb{R}^m$. $(y(\cdot), u(\cdot))$ satisfies (1.1) in the following sense:

$$y(s) = x + \int_t^s f(y(\tau), u(\tau))d\tau, \ \forall\, s \in [t, T].$$

For each $(t, x) \in [0, T] \times \mathbb{R}^n$, we denote $S_{t,x}$ as the set of solutions to (1.1), i.e.

$$S_{t,x} = \{(y(\cdot), u(\cdot)) \,:\, (y(\cdot), u(\cdot)) \text{ satisfies (1.1)}, \ u(\cdot) \in L^\infty(t, T; U)\}.$$

Let $\varphi : \mathbb{R}^n \to \mathbb{R}$ and $\ell : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ be continuous functions. Consider the following optimal control problem: for each $(t, x) \in [0, T] \times \mathbb{R}^n$,

$$\inf \left\{ \int_t^T \ell(y(s), u(s))ds + \varphi(y(T)) \,:\, (y(s), u(s)) \in S_{t,x} \right\}.$$

1.2 The value function

For each $(t, x) \in [0, T] \times \mathbb{R}^n$, we introduce the value function $v : [0, T] \times \mathbb{R}^n \to \mathbb{R}$ defined as

$$v(t, x) = \inf \left\{ \int_t^T \ell(y(s), u(s))ds + \varphi(y(T)) \,:\, (y(\cdot), u(\cdot)) \in S_{t,x} \right\}.$$

We assume the following.

(H1) $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ is continuous and Lipschitz continuous with respect to x uniformly in $u \in U$, i.e. $\exists L_f > 0$ such that $\forall\, x_1, x_2 \in \mathbb{R}^n$, $u \in U$,

$$\|f(x_1, u) - f(x_2, u)\| \leq L_f \|x_1 - x_2\|.$$

(H2) $\ell : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is continuous and Lipschitz continuous with respect to x uniformly in $u \in U$, , i.e. $\exists L_\ell > 0$ such that $\forall\, x_1, x_2 \in \mathbb{R}^n$, $u \in U$,

$$|\ell(x_1, u) - \ell(x_2, u)| \leq L_\ell \|x_1 - x_2\|.$$

(H3) $\varphi : \mathbb{R}^n \to \mathbb{R}$ is Lipschitz continuous, i.e. $\exists L_\varphi > 0$ such that $\forall\, x_1, x_2 \in \mathbb{R}^n$,

$$|\varphi(x_1) - \varphi(x_2)| \leq L_\varphi \|x_1 - x_2\|.$$

Example 1: We take $n = m = 1$, $U = [-1, 1]$, and

$$f(x, u) = u,\ \ \ell \equiv 0,\ \ \varphi(x) = \max\{x, 0\}.$$

Then the value function

$$v(t, x) = \max\{x - T + t, 0\},\ \forall\, t \in [0, T],\ \ x \in \mathbb{R},$$

which is not differentiable.
Example 2: We take the same settings as in Example 1 with φ replaced by

$$\varphi(x) = (\max\{x, 0\})^2\,.$$

Then the value function

$$v(t, x) = (\max\{x - T + t, 0\})^2\,,\ \forall\, t \in [0, T],\ \ x \in \mathbb{R},$$

which is differentiable.
Example 3: We take the same settings as in Example 1 with φ replaced by

$$\varphi(x) = \max\{1 - |x|, 0\}.$$

Then the value function

$$v(t, x) = \max\{1 - |x| - T + t, 0\},\ \forall\, t \in [0, T],\ \ x \in \mathbb{R},$$

which is not differentiable at $x = 0$.
Example 4: We take the same settings as in Example 1 with φ replaced by

$$\varphi(x) = \max\{1 - x^2, 0\},$$

which is differentiable at $x = 0$. Then the value function

$$v(t, x) = \max\{1 - (|x| + T - t)^2, 0\},\ \forall\, t \in [0, T],\ \ x \in \mathbb{R},$$

which is still not differentiable at $x = 0$.

1.3 Dynamic programming principle

v satisfies the following *Dynamic programming principle* (DPP).

Theorem 1. *$\forall\, t \in [0, T]$, $x \in \mathbb{R}^n$, $h \in [0, T - t]$,*

$$v(t, x) = \inf \left\{ v(t + h, y(t + h)) + \int_t^{t+h} \ell(y(s), u(s))ds \,:\, (y(\cdot), u(\cdot)) \in S_{t,x} \right\}.$$

Proof. For any $(y(\cdot), u(\cdot)) \in S_{t,x}$, by the definition of $v(t + h, y(t + h))$ we have for arbitrary $\delta > 0$, there exists $(y^\delta(\cdot), u^\delta(\cdot)) \in S_{t+h,y(t+h)}$ such that

$$v(t + h, y(t + h)) + \delta > \int_{t+h}^T \ell(y^\delta(s), u^\delta(s))ds + \varphi(y^\delta(T)). \tag{1.2}$$

We set

$$\bar{y}(s) = \begin{cases} y(s) & s \in [t, t + h], \\ y^\delta(s) & s \in [t + h, T], \end{cases} \quad \bar{u}(s) = \begin{cases} u(s) & s \in [t, t + h], \\ u^\delta(s) & s \in [t + h, T]. \end{cases}$$

Note that $(\bar{y}(\cdot), \bar{u}(\cdot)) \in S_{t,x}$, we obtain by the definition of $v(t,x)$ that

$$
\begin{aligned}
v(t,x) &\le \int_t^T \ell(\bar{y}(s), u(s))ds + \varphi(\bar{y}(T)) \\
&= \int_t^{t+h} \ell(\bar{y}(s), \bar{u}(s))ds + \int_{t+h}^T \ell(\bar{y}(s), \bar{u}(s))ds + \varphi(\bar{y}(T)) \\
&= \int_t^{t+h} \ell(y(s), u(s))ds + \int_{t+h}^T \ell(y^\delta(s), u^\delta(s))ds + \varphi(y^\delta(T)).
\end{aligned}
$$

Together with (1.2), we deduce that

$$
v(t,x) \le \int_t^{t+h} \ell(y(s), u(s))ds + v(t+h, y(t+h)) + \delta.
$$

The above inequality holds true for arbitrary $\delta > 0$, we have thus

$$
v(t,x) \le \int_t^{t+h} \ell(y(s), u(s))ds + v(t+h, y(t+h)).
$$

Therefore

$$
v(t,x) \le \inf \left\{ \int_h^{t+h} \ell(y(s), u(s))ds + v(t+h, y(t+h)) \,:\, (y(\cdot), u(\cdot)) \in S_{t,x} \right\}.
$$

On the other hand, $\forall \varepsilon > 0$ there exists $(y^\varepsilon(\cdot), u^\varepsilon(\cdot)) \in S_{t,x}$ such that

$$
\begin{aligned}
v(t,x) + \varepsilon &> \int_t^T \ell(y^\varepsilon(s), u^\varepsilon(s))ds + \varphi(y^\varepsilon(T)) \\
&= \int_t^{t+h} \ell(y^\varepsilon(s), u^\varepsilon(s))ds + \int_{t+h}^T \ell(y^\varepsilon(s), u^\varepsilon(s))ds + \varphi(y^\varepsilon(T)) \\
&\ge \int_t^{t+h} \ell(y^\varepsilon(s), u^\varepsilon(s))ds + v(t+h, y^\varepsilon(t+h)) \\
&\ge \inf \left\{ \int_t^{t+h} \ell(y(s), u(s))ds + v(t+h, y(t+h)) \,:\, (y(\cdot), u(\cdot)) \in S_{t,x} \right\}.
\end{aligned}
$$

The above equality holds for arbitrary $\varepsilon > 0$, we have thus

$$
v(t,x) \ge \inf \left\{ \int_t^{t+h} \ell(y(s), u(s))ds + v(t+h, y(t+h)) \,:\, (y(\cdot), u(\cdot)) \in S_{t,x} \right\}.
$$

Finally, we conclude that

$$
v(t,x) = \inf \left\{ \int_t^{t+h} \ell(y(s), u(s))ds + v(t+h, y(t+h)) \,:\, (y(\cdot), u(\cdot)) \in S_{t,x} \right\}.
$$

□

Proposition 2. *Assume (H1)-(H3). The value function v is locally Lipschitz continuous.*

The proof is based on the following two lemmas.

Lemma 3. *Assume (H1)-(H3). For any $t \in [0,T]$, $x_1, x_2 \in \mathbb{R}^n$, there exists $C = C(T, L_f, L_\ell, L_\varphi) > 0$ such that*

$$
|v(t,x_1) - v(t,x_2)| \le C\|x_1 - x_2\|.
$$

Proof. For any $u(\cdot) \in L^\infty(t,T;U)$, let $(y_1(\cdot), u(\cdot)) \in S_{t,x_1}$ and $(y_2(\cdot), u(\cdot)) \in S_{t,x_2}$. We have thus for any $s \in [t,T]$

$$
y_1(s) - y_2(s) = x_1 - x_2 + \int_t^s \left[f(y_1(\tau), u(\tau)) - f(y_2(\tau), u(\tau)) \right] d\tau.
$$

Therefore,

$$\begin{aligned}\|y_1(s)-y_2(s)\| &\leq \|x_1-x_2\| + \int_t^s \|f(y_1(\tau),u(\tau)) - f(y_2(\tau),u(\tau))\|d\tau \\ &\leq \|x_1-x_2\| + L_f \int_t^s \|y_1(\tau)-y_2(\tau)\|d\tau.\end{aligned}$$

By Gronwall's inequality, we deduce that

$$\|y_1(s)-y_2(s)\| \leq e^{L_f(s-t)}\|x_1-x_2\| \leq e^{L_f T}\|x_1-x_2\|.$$

By the definition of $v(t,x_2)$, $\forall\, \varepsilon > 0$ there exists $(y_2^\varepsilon(\cdot), u^\varepsilon(\cdot)) \in S_{t,x_2}$ such that

$$v(t,x_2) + \varepsilon > \int_t^T \ell(y_2^\varepsilon(s), u^\varepsilon(s))ds + \varphi(y_2^\varepsilon(T)).$$

Let $(y_1^\varepsilon(\cdot), u^\varepsilon(\cdot)) \in S_{t,x_1}$, by the definition of $v(t,x_1)$ we have

$$v(t,x_1) \leq \int_t^T \ell(y_1^\varepsilon(s), u^\varepsilon(s))ds + \varphi(y_1^\varepsilon(T)).$$

Therefore,

$$\begin{aligned}v(t,x_1) - v(t,x_2) - \varepsilon &\leq \int_t^T \left[\ell(y_1^\varepsilon(s),u^\varepsilon(s)) - \ell(y_2^\varepsilon(s),u^\varepsilon(s))\right] ds + \varphi(y_1^\varepsilon(T)) - \varphi(y_2^\varepsilon(T)) \\ &\leq L_\ell \int_t^T \|y_1^\varepsilon(s) - y_2^\varepsilon(s)\|ds + L_\varphi \|y_1^\varepsilon(T) - y_2^\varepsilon(T)\| \\ &\leq L_\ell T e^{L_f T}\|x_1-x_2\| + L_\varphi e^{L_f T}\|x_1-x_2\| \\ &= C\|x_1-x_2\|,\end{aligned}$$

where the constant C depends on $T, L_f, L_\ell, L_\varphi$. The above inequality holds true for arbitrary $\varepsilon > 0$, we then deduce that

$$v(t,x_1) - v(t,x_2) \leq C\|x_1-x_2\|.$$

By similar arguments we can obtain that

$$v(t,x_2) - v(t,x_1) \leq C\|x_1-x_2\|,$$

which concludes the proof. □

Lemma 4. *For any $x \in \mathbb{R}^n$, $v(\cdot,x)$ is Lipschitz continuous.*

Proof. For any $x \in \mathbb{R}^n$ and $u \in U$, we have by (H1)

$$\|f(x,u) - f(0,u)\| \leq L_f\|x\|.$$

Consequently,

$$\|f(x,u)\| \leq \|f(0,u)\| + L_f\|x\| \leq M_f + L_f\|x\|,$$

where $M_f = \max_{u\in U}\|f(0,u)\|$. By (H2) and similar arguments we can deduce that

$$\|\ell(x,u)\| \leq \|\ell(0,u)\| + L_\ell\|x\| \leq M_\ell + L_\ell\|x\|,$$

where $M_\ell = \max_{u\in U}\|\ell(0,u)\|$.

Let us take $t_1, t_2 \in [0,T]$ and assume that $t_1 < t_2$ without loss of generality. By applying the dynamic programming principle for $v(t_1,x)$, we have

$$v(t_1,x) = \inf\left\{v(t_2, y(t_2)) + \int_{t_1}^{t_2} \ell(y(s),u(s))ds \,:\, (y(\cdot),u(\cdot)) \in S_{t_1,x}\right\}.$$

For any $(y(\cdot), u(\cdot)) \in S_{t_1,x}$,

$$v(t_1, x) \leq v(t_2, y(t_2)) + \int_{t_1}^{t_2} \ell(y(s), u(s))ds.$$

Therefore,

$$\begin{aligned} v(t_1, x) - v(t_2, x) &\leq v(t_2, y(t_2)) - v(t_2, x) + \int_{t_1}^{t_2} \ell(y(s), u(s))ds \\ &\leq C\|y(t_2) - y(t_1)\| + (M_\ell + L_\ell\|x\|)(t_2 - t_1) \\ &\leq C(M_f + L_f\|x\|)(t_2 - t_1) + (M_\ell + L_\ell\|x\|)(t_2 - t_1) \\ &\leq K(1 + \|x\|)(t_2 - t_1), \end{aligned}$$

where the constant K depends on L_f, L_ℓ, M_f, M_ℓ.

On the other hand for any $\delta > 0$, there exists $(y^\delta(\cdot), u^\delta(\cdot)) \in S_{t_1,x}$ such that

$$v(t_1, x) + \delta \geq v(t_2, y^\delta(t_2)) + \int_{t_1}^{t_2} \ell(y^\delta(s), u^\delta(s))ds.$$

Therefore,

$$\begin{aligned} v(t_1, x) - v(t_2, x) + \delta &\geq v(t_2, y^\delta(t_2)) - v(t_2, x) + \int_{t_1}^{t_2} \ell(y^\delta(s), u^\delta(s))ds \\ &\geq -C\|y^\delta(t_2) - y^\delta(t_1)\| - (M_\ell + L_\ell\|x\|)(t_2 - t_1) \\ &\geq -C(M_f + L_f\|x\|)(t_2 - t_1) - (M_\ell + L_\ell\|x\|)(t_2 - t_1) \\ &\geq -K(1 + \|x\|)(t_2 - t_1). \end{aligned}$$

The above inequality holds true for arbitrary $\delta > 0$, hence

$$v(t_1, x) - v(t_2, x) \geq -K(1 + \|x\|)(t_2 - t_1).$$

We finally conclude that

$$|v(t_1, x) - v(t_2, x)| \leq K(1 + \|x\|)|t_1 - t_2|,$$

which ends the proof. □

2 Lecture 2

2.1 Hamilton-Jacobi-Bellman equation

If the value function is differentiable, we can formally derive the following Hamilton-Jacobi-Bellman equation (HJB) from DPP:

$$\begin{cases} -\partial_t v(t, x) + H(x, D_x v(t, x)) = 0 & \text{in } (0, T) \times \mathbb{R}^n, \\ v(T, x) = \varphi(x) & \text{in } \mathbb{R}^n, \end{cases} \tag{2.1}$$

where $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is called the *Hamiltonian* defined as

$$H(x, p) := \sup_{u \in U} \{-p \cdot f(x, u) - \ell(x, u)\}.$$

However, the value function is not differentiable in general and the notion of solution is unclear for this type of nonlinear hyperbolic equations. In the 1980s, Crandall and Lions have introduced the notion of viscosity solution for nonlinear Hamilton-Jacobi equations. The definition is given as follows.

Definition 5. *Let $u : [0, T] \times \mathbb{R}^n \to \mathbb{R}$ be a continuous function.*

(i) u is called a viscosity supersolution of (2.1) if for any $(t_0, x_0) \in (0, T) \times \mathbb{R}^n$ and $\phi \in C^1((0, T) \times \mathbb{R}^n)$ with $u - \phi$ attains its minimum at (t_0, x_0), we have

$$-\partial_t \phi(t_0, x_0) + H(x_0, D_x \phi(t_0, x_0)) \geq 0,$$

and $u(T, x) \geq \phi(x)$ for any $x \in \mathbb{R}^n$.

(ii) u is called a viscosity subsolution of (2.1) if for any $(t_0, x_0) \in (0,T) \times \mathbb{R}^n$ and $\phi \in C^1((0,T) \times \mathbb{R}^n)$ with $u - \phi$ attains its maximum at (t_0, x_0), we have

$$-\partial_t \phi(t_0, x_0) + H(x_0, D_x \phi(t_0, x_0)) \leq 0,$$

and $u(T, x) \leq \phi(x)$ for any $x \in \mathbb{R}^n$.

(iii) u is called a viscosity solution of (2.1) if u is both a viscosity supersolution and a viscosity subsolution.

Remark 6. *(i) We can assume without loss of generality that $u(t_0, x_0) = \phi(t_0, x_0)$.*

(ii) The minimum/maximum can be replaced by local minimum/maximum.

(iii) The minimum/maximum can be replaced by strict minimum/maximum. Indeed, $\phi(t,x)$ can be replaced by $\phi(t,x) - |t - t_0|^2 - \|x - x_0\|^2 / \phi(t,x) + |t - t_0|^2 + \|x - x_0\|^2$.

(iv) C^1 can be replaced by $C^2, C^3, \ldots, C^\infty$.

(v) The viscosity notion can be defined for Hamilton-Jacobi equations in a general setting, for example

$$-\partial_t u + H(x, u, D_x u, D^2_{xx} u) = 0,$$

and stationary Hamilton-Jacobi equations

$$H(x, u, D_x u, D^2_{xx} u) = 0.$$

One of the motivations of introducing the viscosity notion comes from the vanishing viscosity method for singular perturbation problems. For example, for any $\varepsilon > 0$ consider the following perturbed equation

$$\begin{cases} -\varepsilon u''(x) + |u'(x)| = 1, & -1 < x < 1, \\ u(-1) = u(1) = 0. \end{cases}$$

The equation has a smooth solution

$$u^\varepsilon(x) = 1 - |x| - \varepsilon \left(e^{-\frac{|x|}{\varepsilon}} - e^{-\frac{1}{\varepsilon}} \right), \ -1 \leq x \leq 1.$$

When $\varepsilon \to 0$, u^ε converges to

$$u(x) = 1 - |x|, \ -1 \leq x \leq 1$$

which is a viscosity solution of

$$\begin{cases} |u'(x)| - 1 = 0, & -1 < x < 1, \\ u(-1) = u(1) = 0. \end{cases}$$

In fact, at any $x_0 \neq 0$, $u(\cdot)$ is differentiable and

$$|u'(x_0)| = 1.$$

At $x = 0$, there does not exist $\phi \in C^1(\mathbb{R})$ such that $u - \phi$ attains its minimum at $x = 0$. Therefore u is a viscosity supersolution. On the other hand, for any $\phi \in C^1(\mathbb{R})$ such that $u - \phi$ attains its maximum at $x = 0$, we have

$$u(x) - \phi(x) \leq u(0) - \phi(0), \ \forall -1 < x < 1,$$

i.e.

$$\phi(x) - \phi(0) \geq -|x|, \ \forall -1 < x < 1.$$

By taking $x \to 0_+$ and $x \to 0_-$, we can deduce that

$$|\phi'(0)| \leq 1.$$

Therefore u is a viscosity subsolution. We conclude thus u is a viscosity solution.

Theorem 7. *Assume (H1)-(H3). The value function v is a viscosity solution of (2.1).*

Proof. We first prove that v is a viscosity subsolution. For any $(t_0, x_0) \in (0,T) \times \mathbb{R}^n$ and $\phi \in C^1((0,T) \times \mathbb{R}^n)$ with $u - \phi$ attains its maximum at (t_0, x_0), we have

$$v(t,x) - \phi(t,x) \leq v(t_0, x_0) - \phi(t_0, x_0), \ \forall\, (t,x) \in (0,T) \times \mathbb{R}^n.$$

For any $u \in U$, consider the control function $u(\cdot) \equiv u$ and the associated trajectory $y(\cdot)$ with $(y(\cdot), u(\cdot)) \in S_{t_0, x_0}$, we have by the DPP

$$v(t_0, x_0) \leq v(t_0 + h, y(t_0 + h)) + \int_{t_0}^{t_0+h} \ell(y(s), u) ds, \ \forall\, h \in [0, T - t_0].$$

Note that

$$v(t_0 + h, y(t_0 + h)) - \phi(t_0 + h, y(t_0 + h)) \leq v(t_0, x_0) - \phi(t_0, x_0), \ \forall\, h \in [0, T - t_0],$$

we then deduce that

$$\phi(t_0, x_0) - \phi(t_0 + h, y(t_0 + h)) - \int_{t_0}^{t_0+h} \ell(y(s), u) ds \leq 0, \ \forall\, h \in [0, T - t_0].$$

This is equivalent to

$$\frac{1}{h} \int_{t_0}^{t_0+h} \left[-\partial_t \phi(s, y(s)) - D_x \phi(s, y(s)) \cdot f(y(s), u) - \ell(y(s), u) \right] ds \leq 0, \ \forall\, h \in (0, T - t_0).$$

We obtain by letting $h \to 0$ that

$$-\partial_t \phi(t_0, x_0) - D_x \phi(t_0, x_0) \cdot f(x_0, u) - \ell(x_0, u) \leq 0.$$

The above inequality holds for arbitrary $u \in U$. Consequently,

$$-\partial_t \phi(t_0, x_0) + H(x_0, D_x \phi(t_0, x_0)) \leq 0.$$

On the other hand, we proceed to prove that v is a supersolution. For any $(t_0, x_0) \in (0,T) \times \mathbb{R}^n$ and $\phi \in C^1((0,T) \times \mathbb{R}^n)$ with $u - \phi$ attains its minimum at (t_0, x_0), we have

$$v(t,x) - \phi(t,x) \geq v(t_0, x_0) - \phi(t_0, x_0), \ \forall\, (t,x) \in (0,T) \times \mathbb{R}^n.$$

For any $\varepsilon > 0$, we have by the DPP that there exists $(y^\varepsilon(\cdot), u^\varepsilon(\cdot)) \in S_{t_0, x_0}$ such that

$$v(t_0, x_0) + \varepsilon > v(t_0 + h, y^\varepsilon(t_0 + h)) + \int_{t_0}^{t_0+h} \ell(y^\varepsilon(s), u^\varepsilon(s)) ds \leq 0, \ \forall\, h \in [0, T - t_0].$$

Note that

$$v(t_0 + h, y(t_0 + h)) - \phi(t_0 + h, y(t_0 + h)) \geq v(t_0, x_0) - \phi(t_0, x_0), \ \forall\, h \in [0, T - t_0],$$

we then deduce that

$$\phi(t_0, x_0) - \phi(t_0 + h, y(t_0 + h)) - \int_{t_0}^{t_0+h} \ell(y(s), u) ds + \varepsilon \geq 0, \ \forall\, h \in [0, T - t_0].$$

This is equivalent to

$$\frac{1}{h} \int_{t_0}^{t_0+h} \left[-\partial_t \phi(s, y(s)) - D_x \phi(s, y(s)) \cdot f(y(s), u) - \ell(y(s), u) \right] ds + \sqrt{\varepsilon} \geq 0,$$

where we take $h = \sqrt{\varepsilon}$. By applying the property

$$\frac{1}{h} \int_{t_0}^{t_0+h} g(u(s)) ds \leq \sup_{u \in U} g(u)$$

for any continuous function $g : \mathbb{R}^m \to \mathbb{R}$, and letting $\varepsilon \to 0$, we obtain that

$$-\partial_t \phi(t_0, x_0) + \sup_{u \in U} \{ -D_x \phi(t_0, x_0) \cdot f(x_0, u) - \ell(x_0, u) \} \geq 0,$$

which is

$$-\partial_t \phi(t_0, x_0) + H(x_0, D_x \phi(t_0, x_0)) \geq 0.$$

Finally, v is a viscosity solution since v is both a viscosity subsolution and a viscosity supersolution. □

3 Lecture 3

3.1 Uniqueness result

To simplify the presentation, we consider the case where $\ell \equiv 0$, i.e.

$$v(t,x) = \inf\{\varphi(y(T)) \,:\, (y(\cdot),u(\cdot)) \in S_{t,x}\}.$$

In fact, we can introduce a new state variable

$$z(s) = \int_t^s \ell(y(\tau),u(\tau))d\tau,$$

and the initial optimal control problem with running cost ℓ can be formulated as an equivalent optimal control problem without the running cost. In this case, the HJB equation is thus

$$\begin{cases} -\partial_t v(t,x) + H(x, D_x v(t,x)) = 0 \quad \text{in } (0,T)\times\mathbb{R}^n, \\ v(T,x) = \varphi(x). \end{cases} \tag{3.1}$$

Here the Hamiltonian $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is given by

$$H(x,p) = \sup_{u\in U}\{-f(x,u)\cdot p\}, \;\forall\, x \in \mathbb{R}^n,\; p \in \mathbb{R}^n.$$

H satisfies the following properties. For $x_1, x_2 \in \mathbb{R}^n$ and $p \in \mathbb{R}^n$,

$$|H(x_1,p) - H(x_2,p)| \le L_f\|x_1 - x_2\|\|p\|.$$

For $x \in \mathbb{R}^n$ and $p_1, p_2 \in \mathbb{R}^n$,

$$|H(x,p_1) - H(x,p_2)| \le K(1+\|x\|)\|p_1 - p_2\|,$$

where K depends on M_f and L_f.

Lemma 8. *If $v(t,x)$ is a viscosity solution of (3.1), then for $\lambda > 0$*

$$u(t,x) = e^{\lambda t}v(t,x)$$

is a viscosity solution of

$$\begin{cases} -\partial_t u(t,x) + \lambda u + H(x, D_x u(t,x)) = 0 \quad \textit{in } (0,T)\times\mathbb{R}^n, \\ u(T,x) = e^{\lambda T}\varphi(x). \end{cases} \tag{3.2}$$

We have the following comparison principle result for (3.1).

Theorem 9. *Assume that $u_1, u_2 : [0,T]\times\mathbb{R}^n \to \mathbb{R}$ are bounded and continuous. If u_1 is a viscosity subsolution of (3.2) and u_2 is a viscosity supersolution of (3.2), then*

$$u_1(t,x) \le u_2(t,x),\; \forall\,(t,x) \in [0,T]\times\mathbb{R}^n.$$

Proof. Let

$$M := \sup_{t\in[0,T],x\in\mathbb{R}}\{u_1(t,x) - u_2(t,x)\}.$$

Assume by contradiction that $M > 0$. We construct the following function by the doubling variable technique: for $\varepsilon > 0$, $\beta > 0$ and $0 < m < 1$,

$$\Phi(t,s,x,y) = u_1(t,x) - u_2(s,y) - \frac{|t-s|^2 + \|x-y\|^2}{2\varepsilon} - \beta\left(\langle x\rangle^m + \langle y\rangle^m\right),$$

where

$$\langle x\rangle = \sqrt{1+\|x\|^2},\; \langle y\rangle = \sqrt{1+\|y\|^2}.$$

Note that Φ is continuous and

$$\Phi(t,s,x,y) \to -\infty \text{ as } \max\{\|x\|,\|y\|\} \to \infty,$$

there exists thus $(\bar t, \bar s, \bar x, \bar y)$ such that

$$\sup_{[0,T]^2\times\mathbb{R}^{2n}} \Phi(t,s,x,y) = \Phi(\bar t, \bar s, \bar x, \bar y).$$

By the definition of M, there exists $(\tilde t, \tilde x)$ such that

$$u_1(\tilde t, \tilde x) - u_2(\tilde t, \tilde x) > \frac{M}{2}.$$

Therefore,

$$\Phi(\bar t, \bar s, \bar x, \bar y) \ge \Phi(\tilde t, \tilde t, \tilde x, \tilde x) = u_1(\tilde t, \tilde x) - u_2(\tilde t, \tilde x) - \beta\left(\langle \tilde x\rangle^m + \langle \tilde y\rangle^m\right) > \frac{M}{2} - \beta\left(\langle \tilde x\rangle^m + \langle \tilde y\rangle^m\right).$$

We then choose $\beta > 0$ such that

$$\Phi(\bar t, \bar s, \bar x, \bar y) > \frac{M}{4},$$

which implies that

$$\beta\left(\langle \bar x\rangle^m + \langle \bar y\rangle^m\right) \le u_1(\bar t, \bar x) - u_2(\bar s, \bar y) - \frac{M}{4}.$$

Since u_1 and u_2 are bounded, there exists $C_1 > 0$ such that

$$\beta\left(\langle \bar x\rangle^m + \langle \bar y\rangle^m\right) \le C_1.$$

Hence,

$$\|\bar x\|, \|\bar y\| \in \Omega := \overline{B\left(0, \left(\frac{C_1}{\beta}\right)^{1/m}\right)}.$$

Note that u_1 and u_2 are uniformly continuous in $[0,T]\times\Omega$, there exists a modulus of continuity $\omega(\cdot)$ such that

$$|u_i(t,x) - u_i(s,y)| \le \omega\left(|t-s| + \|x-y\|\right),\ i=1,2,\ t,s\in[0,T],\ x,y\in\Omega.$$

By the fact that $\Phi(\bar t, \bar t, \bar x, \bar x) \le \Phi(\bar t, \bar s, \bar x, \bar y)$ and $\Phi(\bar s, \bar s, \bar y, \bar y) \le \Phi(\bar t, \bar s, \bar x, \bar y)$, we have

$$\Phi(\bar t, \bar t, \bar x, \bar x) + \Phi(\bar s, \bar s, \bar y, \bar y) \le 2\Phi(\bar t, \bar s, \bar x, \bar y),$$

where we deduce that

$$\frac{|\bar t - \bar s|^2 + \|\bar x - \bar y\|^2}{\varepsilon} \le u_1(\bar t, \bar x) - u_1(\bar s, \bar y) + u_2(\bar t, \bar x) - u_2(\bar s, \bar y).$$

Since u_1 and u_2 are bounded and uniformly continuous in $[0,T]\times\Omega$, we obtain

$$\frac{|\bar t - \bar s|^2 + \|\bar x - \bar y\|^2}{\varepsilon} \le C,\ \ \frac{|\bar t - \bar s|^2 + \|\bar x - \bar y\|^2}{\varepsilon} \le 2\omega\left(|\bar t - \bar s| + \|\bar x - \bar y\|\right).$$

The above inequalities imply that

$$|\bar t - \bar s| \le \sqrt{C\varepsilon},\ \ \|\bar x - \bar y\| \le \sqrt{C\varepsilon},\ \ \frac{|\bar t - \bar s|^2 + \|\bar x - \bar y\|^2}{\varepsilon} \le 2\omega\left(2\sqrt{C\varepsilon}\right).$$

For any $(t,x)\in[0,T]\times\mathbb{R}^n$, we set

$$\phi_1(t,x) = u_2(\bar s, \bar y) + \frac{|t - \bar s|^2 + \|x - \bar y\|^2}{2\varepsilon} + \beta\left(\langle x\rangle^m + \langle \bar y\rangle^m\right).$$

Hence $u_1(t,x) - \phi_1(t,x)$ attains its maximum at $\bar t, \bar x$. Since u_1 is a viscosity subsolution of (3.2), we have thus

$$-\partial_t\phi_1(\bar t, \bar x) + \lambda u_1(\bar t, \bar x) + H(\bar x, D_x\phi_1(\bar t, \bar x)) \le 0.$$

That is

$$-\frac{\bar t - \bar s}{\varepsilon} + \lambda u_1(\bar t, \bar x) + H\left(\bar x, \frac{\bar x - \bar y}{\varepsilon} + \beta m\langle \bar x\rangle^{m-2}\bar x\right) \le 0. \tag{3.3}$$

For any $(s,y)\in[0,T]\times\mathbb{R}^n$, we set

$$\phi_2(s,y)=u_1(\bar{t},\bar{x})-\frac{|\bar{t}-s|^2+\|\bar{x}-y\|^2}{2\varepsilon}-\beta\left(\langle\bar{x}\rangle^m+\langle y\rangle^m\right).$$

Hence $u_2(s,y)-\phi_2(s,y)$ attains its minimum at $\bar{s},\bar{y}$. Since u_2 is a viscosity supersolution of (3.2), we have thus

$$-\partial_s\phi_2(\bar{s},\bar{y})+\lambda u_2(\bar{s},\bar{y})+H(\bar{y},D_y\phi_2(\bar{s},\bar{y}))\geq 0.$$

That is

$$-\frac{\bar{t}-\bar{s}}{\varepsilon}+\lambda u_2(\bar{s},\bar{y})+H\left(\bar{y},\frac{\bar{x}-\bar{y}}{\varepsilon}-\beta m\langle\bar{y}\rangle^{m-2}\bar{y}\right)\geq 0. \tag{3.4}$$

The substraction of (3.3) and (3.4) gives then

$$\lambda[u_1(\bar{t},\bar{x})-u_2(\bar{s},\bar{y})]\leq H\left(\bar{y},\frac{\bar{x}-\bar{y}}{\varepsilon}-\beta m\langle\bar{y}\rangle^{m-2}\bar{y}\right)-H\left(\bar{x},\frac{\bar{x}-\bar{y}}{\varepsilon}+\beta m\langle\bar{x}\rangle^{m-2}\bar{x}\right).$$

The properties of the Hamiltonian H imply that

$$\begin{aligned}
& H\left(\bar{y},\frac{\bar{x}-\bar{y}}{\varepsilon}-\beta m\langle\bar{y}\rangle^{m-2}\bar{y}\right)-H\left(\bar{x},\frac{\bar{x}-\bar{y}}{\varepsilon}+\beta m\langle\bar{x}\rangle^{m-2}\bar{x}\right)\\
\leq\;& L_f\|\bar{x}-\bar{y}\|\left\|\frac{\bar{x}-\bar{y}}{\varepsilon}-\beta m\langle\bar{y}\rangle^{m-2}\bar{y}\right\|+K(1+\|x\|)\cdot 2\beta m\left(\langle\bar{x}\rangle^{m-2}\|\bar{x}\|+\langle\bar{y}\rangle^{m-2}\|\bar{y}\|\right)\\
\leq\;& L_f\frac{\|\bar{x}-\bar{y}\|^2}{\varepsilon}+\beta m(L_f+2K)\|\bar{x}-\bar{y}\|+4\beta mK+2\beta mK\left(\langle\bar{x}\rangle^m+\langle\bar{y}\rangle^m\right)\\
\leq\;& 2L_f\omega\left(2\sqrt{C\varepsilon}\right)+\beta m(L_f+2K)\sqrt{C\varepsilon}+4\beta mK+2mC_1K.
\end{aligned}$$

Together with the property $u_1(\bar{t},\bar{x})-u_2(\bar{s},\bar{y})>\frac{M}{4}$, we deduce that

$$\frac{\lambda M}{4}\leq 2L_f\omega\left(2\sqrt{C\varepsilon}\right)+\beta m(L_f+2K)\sqrt{C\varepsilon}+4\beta mK+2mC_1K.$$

By letting $\beta\to 0$ and $\varepsilon\to 0$, we obtain

$$\frac{\lambda M}{4}\leq 2mC_1K. \tag{3.5}$$

Note that M,C_1,K are independent from the choice of ε and β, and m can be any number in $(0,1)$. We can choose

$$m<\min\left\{\frac{\lambda M}{8C_1K},1\right\}$$

which is a contradiction to (3.5). Therefore $M\leq 0$ which ends the proof. □

Remark 10. *The term $\frac{|t-s|^2+\|x-y\|^2}{2\varepsilon}$ serves as a regularization of u_1 and u_2. Indeed, for any $w\in C(\mathbb{R}^d)$, consider*

$$w_\varepsilon(x):=\inf_{y\in\mathbb{R}^d}\left\{w(y)+\frac{\|x-y\|^2}{\varepsilon}\right\},\ w^\varepsilon(x):=\sup_{y\in\mathbb{R}^d}\left\{w(y)-\frac{\|x-y\|^2}{\varepsilon}\right\},$$

which are called inf-convolution and sup-convolution of w respectively. Let us take $w(x)=|x|,\ x\in\mathbb{R}$ for example. Then by direct computation,

$$w_\varepsilon(x)=\begin{cases}-x-\frac{\varepsilon}{4} & x<-\frac{\varepsilon}{2},\\ \frac{x^2}{\varepsilon} & -\frac{\varepsilon}{2}\leq x\leq\frac{\varepsilon}{2},\\ x-\frac{\varepsilon}{4} & x>\frac{\varepsilon}{2}.\end{cases}$$

Here w_ε is a smooth function and $w_\varepsilon\to w$ pointwise when $\varepsilon\to 0$.

Corollary 11. *(3.2) has a unique viscosity solution, and so does (3.1).*

4 Lecture 4

4.1 Invariance results

Theorem 12. *Let $w \in C([0,T] \times \mathbb{R}^n)$.*

(i) w satisfies the super-optimality, i.e. $\forall\, \varepsilon > 0$, there exists $(y^\varepsilon(\cdot), u^\varepsilon(\cdot)) \in S_{t,x}$ such that

$$w(t,x) + \varepsilon \geq w(t+h, y(t+h)),\ \forall\, h \in [0, T-t]$$

and $w(T,x) \geq \varphi(x)$ if and only if w is a viscosity supersolution of (2.1).

(ii) w satisfies the sub-optimality, i.e. for any $(y^\varepsilon(\cdot), u(\cdot)) \in S_{t,x}$ such that

$$w(t,x) \leq w(t+h, y(t+h)),\ \forall\, h \in [0, T-t]$$

and $w(T,x) \leq \varphi(x)$ if and only if w is a viscosity subsolution of (2.1).

The proof and more details on the invariance properties refer to [4]. Based on the above result, here we provide another proof for the comparison principle of viscosity solution. Let $v_1, v_2 \in C([0,T] \times \mathbb{R}^n)$ be a viscosity subsolution and supersolution respectively. Then $\forall\, \varepsilon > 0$, there exists $(y^\varepsilon(\cdot), u^\varepsilon(\cdot)) \in S_{t,x}$ such that

$$v_2(t,x) + \varepsilon \geq v_2(T, y(T)) \geq \varphi(x).$$

Besides,

$$v_1(t,x) \leq v_1(T, y(T)) \leq \varphi(x).$$

Therefore

$$v_1(t,x) \leq v_2(t,x) + \varepsilon,\ \forall\, \varepsilon > 0.$$

We then conclude that

$$v_1(t,x) \leq v_2(t,x).$$

4.2 Stability

Before we introduce the stability result, we give the following technical lemma.

Lemma 13. *Let $\Omega \subset \mathbb{R}^d$ be an open set, and $w_n, w \in C(\Omega)$ with*

$$w_n(\cdot) \to w(\cdot) \text{ in } C(\Omega).$$

If $x_0 \in \Omega$ is a strict local maximum of w, then there exists a sequence of local maximum of w_n, denoted as $(x_n)_{n\in\mathbb{N}}$, such that $x_n \to x_0$ as $n \to \infty$.

The stability result of viscosity solution is given as follows.

Theorem 14. *$\forall\, \varepsilon > 0$, $u^\varepsilon \in C([0,T] \times \mathbb{R}^n)$ is a viscosity subsolution/supersolution of*

$$\begin{cases} -\partial_t u^\varepsilon(t,x) + H_\varepsilon(x, D_x u^\varepsilon(t,x)) = 0 & \text{in } (0,T) \times \mathbb{R}^n, \\ u^\varepsilon(T,x) = \varphi^\varepsilon(x). \end{cases}$$

If $u^\varepsilon \to u$ in $C([0,T] \times \mathbb{R}^n)$, $H_\varepsilon \to H$ in $C(\mathbb{R}^n \times \mathbb{R}^n)$ and $\varphi^\varepsilon \to \varphi$ in $C(\mathbb{R}^n)$, then u is a viscosity subsolution/supersolution of

$$\begin{cases} -\partial_t u(t,x) + H(x, D_x u(t,x)) = 0 & \text{in } (0,T) \times \mathbb{R}^n, \\ u(T,x) = \varphi(x). \end{cases}$$

Proof. For any $(t_0, x_0) \in (0,T) \times \mathbb{R}^n$, $\phi \in C^1((0,T) \times \mathbb{R}^n)$ with $u - \phi$ attains its maximum at (t_0, x_0). Therefore,

$$u - \phi - (t - t_0)^2 - \|x - x_0\|^2$$

attains its strict maximum at (t_0, x_0). Hence, $u^\varepsilon - \phi - (t-t_0)^2 - \|x - x_0\|^2$ attains its local maximum at $(t_\varepsilon, x_\varepsilon)$ and $(t_\varepsilon, x_\varepsilon) \to (t_0, x_0)$ as $\varepsilon \to 0$. Since u^ε is a viscosity subsolution, we have thus

$$-\partial_t \phi(t_\varepsilon, x_\varepsilon) - 2(t_\varepsilon - t_0) + H_\varepsilon(x_\varepsilon, D_x \phi(t_\varepsilon, x_\varepsilon) + 2\|x_\varepsilon - x_0\|) \leq 0.$$

By letting $\varepsilon \to 0$ and the local uniform convergence of H_ε, we obtain

$$-\partial_t \phi(t_0, x_0) + H(x_0, D_x \phi(t_0, x_0)) \leq 0,$$

which implies that u is a viscosity subsolution. ∎

5 Lecture 5

In this section, we discuss the numerical methods for the computation of the viscosity solutions of HJB equations as follows:

$$\begin{cases} -\partial_t u(t,x) + H(x, D_x u(t,x)) = 0 & \text{in } (0,T)\times\mathbb{R}^d, \\ u(T,x) = \varphi(x). \end{cases}$$

The above PDE with a final condition can be turned into an equivalent PDE with an initial condition by introducing

$$w(t,x) = v(T-t,x),\ \forall\, t\in[0,T],\ x\in\mathbb{R}^d.$$

Then w is the viscosity solution of

$$\begin{cases} \partial_t w(t,x) + H(x, D_x w(t,x)) = 0 & \text{in } (0,T)\times\mathbb{R}^d, \\ w(0,x) = \varphi(x). \end{cases}$$

Besides, w is also the value function of the following optimal control problem

$$w(t,x) = \inf\left\{\int_0^t \ell(y(s),u(s))ds + \varphi(y(t))\ :\ (y(\cdot),u(\cdot))\in S_{0,x}\right\}.$$

The Dynamic Programming Principle for w is the following: for any $t\in[0,T]$, $x\in\mathbb{R}^d$ and $h\in[0,t]$.

$$w(t,x) = \inf\left\{w(t-h,y(h)) + \int_0^h \ell(y(s),u(s))ds\ :\ (y(\cdot),u(\cdot))\in S_{0,x}\right\}.$$

To simply the presentation, we take $\ell\equiv 0$ in the sequel.

5.1 Finite difference schemes

Consider the following discretization: let $N_T\in\mathbb{N}$ and $\Delta x>0$, we set

$$\Delta t = \frac{T}{N_T},\ t_n = n\Delta t,\ \text{for } n=0,1,\ldots,N_T,$$

and

$$x_I = I\Delta x,\ \forall\, I\in\mathbb{Z}^d.$$

We denote by W_I^n as the numerical approximation of $w(t_n,x_I)$. For each $I=(i_1,i_2,\ldots,i_d)\in\mathbb{Z}^d$ and $j=1,2,\ldots,d$, we set

$$D_j^+ W_I^n = \frac{W^n_{(i_1,\ldots,i_j+1,\ldots,i_d)} - W_I^n}{\Delta x},\ \ D_j^- W_I^n = \frac{W_I^n - W^n_{(i_1,\ldots,i_j-1,\ldots,i_d)}}{\Delta x}.$$

Then the finite difference scheme is the following:

- $W_I^0 = \varphi(x_I)$, $\forall\, I\in\mathbb{Z}^d$.
- For $n=0,\ldots,N_T-1$,

$$\frac{W_I^{n+1}-W_I^n}{\Delta t} = \min_{u\in U}\left\{\max_{j=1,2,\ldots,d}\left\{\max\{f_j(x_I,u),0\}\cdot D_j^+W_I^n + \min\{f_j(x_I,u),0\}\cdot D_j^-W_I^n\right\}\right\}.$$

Proposition 15. *Under the CFL condition*

$$\max_{x\in\mathbb{R}^d,u\in U}\{\|f(x,u)\|\}\frac{\Delta t}{\Delta x}\leq 1,$$

the finite difference scheme is monotone, i.e. for two initial conditions φ and $\tilde\varphi$ with $\varphi\leq\tilde\varphi$, the associated numerical solutions W_I^n and $\tilde W_I^n$ satisfies

$$W_I^n \leq \tilde W_I^n,\ \forall\, n=0,1,\ldots,N_T,\ I\in\mathbb{Z}^d.$$

Under the monotonicity property of the scheme, we can prove that the numerical solution converges to the viscosity solution and the convergence rate of this type of scheme can refer to [1].

5.2 Semi-Lagrangian schemes

The semi-Lagrangian scheme is based on the discretization of the DPP: for any $t \in [0, T]$, $x \in \mathbb{R}^d$ and $h \in [0, t]$.

$$w(t, x) = \inf \{ w(t-h, y(h)) \,:\, (y(\cdot), u(\cdot)) \in S_{0,x} \}.$$

We take $h = \Delta t$ and the scheme is the following:

- $W_I^0 = \varphi(x_I)$, $\forall I \in \mathbb{Z}^d$.
- For $n = 0, \ldots, N_T - 1$,
$$W_I^{n+1} = \min_{u \in U} Interp[W^n](x_I + f(x_I, u)\Delta t).$$

Here $Interp[W^n](x)$ denotes the interpolation of the function $w(t_n, \cdot)$ at the point x. The following result gives the error estimate between $y(\Delta t)$ and its linear approximation $x_I + f(x_I, u)\Delta t$.

Proposition 16. *Assume (H1), f is bounded by some constant $M_f > 0$ and that $\{f(x, u) \,:\, u \in U\}$ is convex for any $x \in \mathbb{R}^d$. For any $(y(\cdot), u(\cdot)) \in S_{0,x}$, there exists $u \in U$ such that*

$$y(\Delta t) = x + f(x, u)\Delta t + O(\Delta t^2).$$

Proof. For any $s \in [0, \Delta t]$, we have

$$\|y(s) - x\| \leq M_f s.$$

We define the set-valued function

$$F(x) := \{f(x, u) \,:\, u \in U\}, \ \forall x \in \mathbb{R}^d.$$

By (H2), we have for any $x_1, x_2 \in \mathbb{R}^d$

$$F(x_1) \subseteq F(x_2) + B(0, L_f \|x_1 - x_2\|).$$

Therefore,

$$\begin{aligned}
y(\Delta t) - x &= \int_0^{\Delta t} f(y(s), u(s)) ds \\
&\subseteq \int_0^{\Delta t} F(y(s)) ds \\
&\subseteq \int_0^{\Delta t} [F(x) + B(0, L_f \|y(s) - x\|)] \, ds \\
&\subseteq \Delta t F(x) + \int_0^{\Delta t} B(0, L_f M_f s) ds \\
&\subseteq \Delta t F(x) + B(0, \frac{1}{2} L_f M_f \Delta t^2).
\end{aligned}$$

Hence there exists $u \in U$ such that

$$\|y(\Delta t) - x - f(x, u)\Delta t\| \leq \frac{1}{2} L_f M_f \Delta t^2.$$

□

Suppose that the interpolation in the semi-Lagrangian scheme is first-order. Then we have the following error estimation result.

Theorem 17. *Assume that w is Lipschitz continuous with the Lipschitz constant $L_w > 0$. Then there exists a constant $C = C(L_u, L_f, M_f, T)$ such that*

$$|W_I^n - w(t_n, x_I)| \leq C \left(\frac{\Delta x}{\Delta t} + \Delta t \right), \ \forall n = 0, 1, \ldots, N_T, \ I \in \mathbb{R}^d.$$

Proof. For any $n = 0, 1, \ldots, N_T - 1,\ I \in \mathbb{R}^d$,

$$
\begin{aligned}
& |W_I^{n+1} - w(t_{n+1}, x_I)| \\
\leq\ & \left| \min_{u \in U} Interp[W^n](x_I + f(x_I, u)\Delta t) - \min_{(y(\cdot), u(\cdot)) \in S_{0,x}} w(t_n, y(\Delta t)) \right| \\
\leq\ & \left| \min_{u \in U} Interp[W^n](x_I + f(x_I, u)\Delta t) - \min_{u \in U} w(t_n, x_I + f(x_I, u)\Delta t) \right| + \frac{1}{2} L_w L_f M_f \Delta t^2 \\
\leq\ & \max_{u \in U} |Interp[W^n](x_I + f(x_I, u)\Delta t) - w(t_n, x_I + f(x_I, u)\Delta t)| + \frac{1}{2} L_w L_f M_f \Delta t^2 \\
\leq\ & \max_{x \in \mathbb{R}^d} |Interp[W^n](x) - w(t_n, x)| + \frac{1}{2} L_w L_f M_f \Delta t^2 \\
\leq\ & \max_{I \in \mathbb{Z}^d} |Interp[W^n](x_I) - w(t_n, x_I)| + L_w \sqrt{d} \Delta x + \frac{1}{2} L_w L_f M_f \Delta t^2 \\
=\ & \max_{I \in \mathbb{Z}^d} |W_I^n - w(t_n, x_I)| + L_w \sqrt{d} \Delta x + \frac{1}{2} L_w L_f M_f \Delta t^2 \\
\leq\ & \max_{I \in \mathbb{Z}^d} \left| W_I^0 - w(t_0, x_I) \right| + N_T L_w \sqrt{d} \Delta x + \frac{1}{2} N_T L_w L_f M_f \Delta t^2 \\
=\ & C \left(\frac{\Delta x}{\Delta t} + \Delta t \right),
\end{aligned}
$$

where $C = T L_w \sqrt{d} + \frac{1}{2} T L_w L_f M_f$. □

References

[1] G. Barles and E. Jakobsen, *On the convergence rate of approximation schemes for Hamilton-Jacobi-Bellman equations*, ESAIM: M2AN, Vol. 36, No 1, 2002, pp. 33-54. DOI: 10.1051/m2an:2002002.

[2] M. Bardi and I. Capuzzo-Dolcetta, *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations*, Systems and Control: Foundations and Applications, Birkhäuser, Boston, 1997.

[3] P. Cannarsa and C. Sinestrari, *Semiconcave Functions, Hamilton-Jacobi Equations, and Optimal Control*, Progress in Nonlinear Differential Equations and their Applications, 58. Birkhäuser Boston, Inc., Boston, MA 2004.

[4] F.H. Clarke, Yu.S. Ledyaev, R.J. Stern and P.R. Wolenski, *Nonsmooth Analysis and Control Theory*, Graduate Texts in Mathematics 178, Springer-Verlag, New York.

[5] M. Falcone and R. Ferretti, *Semi-Lagrangian Approximation Schemes for Linear and Hamilton-Jacobi Equations*, Society for Industrial and Applied Mathematics, 2013.

[6] J.M. Yong and H.W. Lou, *A concise introduction to optimal control theory*, Higher Education Press, 2006.