Chapter 8

Dynamic programming

8.1 Dynamic programming principle

We can now state the DPP.

Theorem 8.1.1 (Dynamic programming principle) *The value function v satisifies*

$$v(t,x) \quad = \quad \sup_{\alpha\in\mathcal{A}_t} \mathbb{E}\Big[v\big(\theta^\alpha, X^{t,x,\alpha}_{\theta^\alpha}\big) + \int_t^{\theta^\alpha} f(X^{t,x,\alpha}_s,\alpha_s)ds\Big]$$

for any $(t,x)\in[0,T]\times\mathbb{R}^n$ and any family of stopping times $\{\theta^\alpha,\ \alpha\in\mathcal{A}\}$ valued in $[t,T]$.

Proof. To alleviate notations, we omit the dependence of θ in α. We proceed in two steps by proving that each term of the equality is greater than the other.

Step 1. Fix $(t,x)\in[0,T]\times\mathbb{R}^n$ and $\alpha\in\mathcal{A}_t$. By the conditioning property for controlled diffusions, we can find for $\mathbb{P}$-a.a. $\omega\in\Omega$ a control $\tilde\alpha^\omega$ such that

$$\begin{aligned}
&\mathbb{E}\Big[g(X^{t,x,\alpha}_T) + \int_t^T f(X^{t,x,\alpha}_s,\alpha_s)ds\Big|\mathcal{F}_\theta\Big](\omega)\\
= \quad & J(\theta(\omega), X^{t,x,\alpha}_\theta(\omega),\alpha^\omega) + \int_t^{\theta(\omega)} f(X^{t,x,\alpha}_s,\alpha_s)ds\ .
\end{aligned}$$

Since $J\leq v$, we get from the tower property

$$v(t,x) \quad \leq \quad \sup_{\alpha\in\mathcal{A}_t} \mathbb{E}\Big[v(\theta, X^{t,x,\alpha}_\theta) + \int_t^\theta f(X^{t,x,\alpha}_s,\alpha_s)ds\Big]\ .$$

Step 2. We now prove the reverse inequality. Fix $\alpha \in \mathcal{A}$, $(t_0, x_0) \in [0,T) \times \mathbb{R}^n$, a ball $B_0 = B((t_0,x_0), r)$ of radius r centered on (t_0, x_0) and fix a compact subset Θ of $[0,T] \times \mathbb{R}^n$ such that $B_0 \subset \Theta$. Let $(B_n)_{n\geq 1}$ be a partition of Θ and $(t_n, x_n)_{n\geq 1}$ be a sequence such that $(t_n, x_n) \in B_n$ for each $n \geq 1$. By definition, for each $n \geq 1$, we can find $\alpha^n \in \mathcal{A}$ such that

$$J(t_n, x_n, \alpha^n) \;\; \geq \;\; v(t_n, x_n) - \varepsilon \qquad (8.1.1)$$

with $\varepsilon > 0$. Moreover, from the local uniform continuity of $J(., \alpha)$ and v, we can chose $(t_n, x_n, B_n)_{n\geq 1}$ such that

$$B_n \;\; \subset \;\; [t_n - \eta, t_n] \times B(x_n, \eta)$$

for some $\eta > 0$ and

$$|v(.) - v(t_n, x_n)| + |J(., \alpha_n) - J(t_n, x_n, \alpha_n)| \;\; \leq \;\; \varepsilon \text{ on } B_n\,. \qquad (8.1.2)$$

Let us now define the stopping time

$$\vartheta \;\; = \;\; \inf\{s \in [t_0, T] \;:\; (s, X^{t_0,x_0,\alpha}) \notin B_0\} \wedge \theta$$

where θ is a given stopping time valued in $[t_0, T]$. We next define the control $\bar{\alpha}$ by

$$\bar{\alpha}_t \;\; = \;\; \alpha_t \mathbb{1}_{t<\vartheta} + \mathbb{1}_{t\geq\vartheta}\Big(\sum_{n\geq 1} \alpha_t^n \mathbb{1}_{(\vartheta, X_\vartheta^{t_0,x_0,\alpha}) \in B_n}\Big)\,, \quad t \geq 0\,.$$

Since α^n is independent of $\mathcal{F}_{t_n}$ for each $n \geq 1$, we get from (8.1.1) and (8.1.2)

$$\begin{aligned}
J(t_0, x_0, \bar{\alpha}) \;\; & \geq \;\; \mathbb{E}\Big[J(\vartheta, X_\vartheta^{t_0,x_0,\bar{\alpha}}, \bar{\alpha}) + \int_{t_0}^{\vartheta} f(X_s^{t_0,x_0,\bar{\alpha}}, \bar{\alpha}_s) ds\Big] \\
& \geq \;\; \mathbb{E}\Bigg[\sum_{n\geq 1}\Big(J(t_n, x_n, \alpha^n) - \varepsilon + \int_{t_0}^{\vartheta} f(X_s^{t_0,x_0,\alpha}, \alpha_s) ds\Big) \mathbb{1}_{(\vartheta, X_\vartheta^{t_0,x_0,\alpha}) \in B_n}\Bigg] \\
& \geq \;\; \mathbb{E}\Bigg[\sum_{n\geq 1}\Big(v(t_n, x_n) - 2\varepsilon + \int_{t_0}^{\vartheta} f(X_s^{t_0,x_0,\alpha}, \alpha_s) ds\Big) \mathbb{1}_{(\vartheta, X_\vartheta^{t_0,x_0,\alpha}) \in B_n}\Bigg] \\
& \geq \;\; \mathbb{E}\Big[v(\vartheta, X_\vartheta^{t_0,x_0,\alpha}) + \int_{t_0}^{\vartheta} f(X_s^{t_0,x_0,\alpha}, \alpha_s) ds\Big] - 3\varepsilon\,.
\end{aligned}$$

Therefore, we get

$$v(t_0, x_0) \;\; \geq \;\; \mathbb{E}\Big[v(\vartheta, X_\vartheta^{t_0,x_0,\alpha}) + \int_{t_0}^{\vartheta} f(X_s^{t_0,x_0,\alpha}, \alpha_s) ds\Big]\,.$$

Letting $r \to +\infty$, we get $\vartheta \to \theta$ and by dominated convergence, we get

$$v(t_0, x_0) \;\; \geq \;\; \mathbb{E}\Big[v(\theta, X_\theta^{t_0,x_0,\alpha}) + \int_{t_0}^{\theta} f(X_s^{t_0,x_0,\alpha}, \alpha_s) ds\Big]\,.$$

8.2 HJB equation

We prove in this section that, if v is smooth enough, it solves a PDE called the Hamilton Jacobi Bellman (HJB) equation. More precisely, define the second order local operator $\mathcal{L}^a$, for $a \in A$ by

$$\mathcal{L}^a\varphi(t,x) \quad = \quad \partial_t\varphi(t,x) + b(t,x,a).\nabla\varphi(t,x) + \frac{1}{2}\mathrm{Tr}\big(\sigma(t,x,a)\sigma^\top(t,x,a)\nabla^2\varphi(t,x)\big)$$

for $(t,x) \in [0,T) \times \mathbb{R}^n$ and $\varphi \in C^{1,2}([0,T) \times \mathbb{R}^n)$. Then, the HJB equation takes the following form

$$\sup_{a\in A}\big\{\mathcal{L}^a v(t,x) + f(x,a)\big\} \quad = \quad 0 \;, \quad (t,x) \in [0,T) \times \mathbb{R}^n \tag{8.2.1}$$

together with the terminal condition

$$v(T,x) \quad = \quad g(x) \;, \quad x \in \mathbb{R}^n \;. \tag{8.2.2}$$

To simplify notations, we denote by $\mathcal{H}$ the operator defined by

$$\mathcal{H}\varphi(t,x) \quad = \quad \sup_{a\in A}\big\{\mathcal{L}^a\varphi(t,x) + f(x,a)\big\}$$

for $(t,x) \in [0,T) \times \mathbb{R}^n$ and $\varphi \in C^{1,2}([0,T) \times \mathbb{R}^n)$.

Theorem 8.2.1 *Suppose that $v \in C([0,T] \times \mathbb{R}^n) \cap C^{1,2}([0,T) \times \mathbb{R}^n)$. Then, v is a solution to* (8.2.1)-(8.2.2).

Proof. Fix $(t,x) \in [0,T) \times \mathbb{R}^d$ and assume that $\mathcal{H}v(t,x) \neq 0$. We then distinguish two cases and work toward a contradiction in each of them.

Case 1: $\mathcal{H}v(t,x) > 0$. Let $a \in A$ such that

$$\mathcal{L}^a\varphi(t,x) + f(x,a) \quad > \quad 0 \;.$$

By continuity of the involved functions, there exists a compact neighborhood $V \subset [0,T) \times \mathbb{R}^n$ of (t,x) such that

$$\mathcal{L}^a\varphi(.) + f(.,a) \quad > \quad 0 \text{ on } V \;. \tag{8.2.3}$$

Let $\alpha = a$, be the constant control of $\mathcal{A}$ equal to a, and θ be the first exit time of $(s, X_s^{t,x,\alpha})$ from V. From Itô's formula we have

$$\begin{aligned} & \mathbb{E}\Big[v\big(\theta, X_\theta^{t,x,\alpha}\big) + \int_t^\theta f(X_s^{t,x,\alpha}, \alpha_s)ds\Big] \\ = \;\; & v(t,x) + \mathbb{E}\Big[\int_t^\theta \big(\mathcal{L}^a v(s, X_s^{t,x,\alpha}) + f(X_s^{t,x,\alpha}, a)\big)ds\Big] \ . \end{aligned}$$

In view of (8.2.3) and since $\theta > 0$ a.s., this contradict the DPP.

Case 2: $\mathcal{H}v(t,x) < 0$. Still using the continuity of the involved functions, this implies

$$\mathcal{H}v \;\; < \;\; 0$$

on $V := B((t,x),r) \subset [0,T) \times \mathbb{R}^n$. Moreovero, for r small enough, we also have

$$\mathcal{H}w \;\; \leq \;\; 0 \text{ on } V \tag{8.2.4}$$

where the function w is defined by

$$w(s,y) \;\; = \;\; v(s,y) + (s-t) + |y-x|^2 \ , \quad (s,y) \in [0,T) \times \mathbb{R}^n \ .$$

For $\alpha \in \mathcal{A}_t$, let θ be the first exit time of $(s, X^{t,x,\alpha})$ from V. Using Itô's formula and (8.2.4), we have

$$\begin{aligned} & \mathbb{E}\Big[w\big(\theta, X_\theta^{t,x,\alpha}\big) + \int_t^\theta f(X_s^{t,x,\alpha}, \alpha_s)ds\Big] \\ = \;\; & w(t,x) + \mathbb{E}\Big[\int_t^\theta \big(\mathcal{L}^a w(s, X_s^{t,x,\alpha}) + f(X_s^{t,x,\alpha}, a)\big)ds\Big] \\ \leq \;\; & v(t,x) \ . \end{aligned}$$

From the definition of w and θ, it follows that

$$\begin{aligned} v(t,x) \;\; \geq \;\; & \mathbb{E}\Big[(\theta - t) + |X_\theta^{t,x,\alpha} - x|^2\Big] \\ & + \mathbb{E}\Big[v\big(\theta, X_\theta^{t,x,\alpha}\big) + \int_t^\theta f(X_s^{t,x,\alpha}, \alpha_s)ds\Big] \ . \end{aligned}$$

We then notice from the definition of θ that

$$(\theta - t) + |X^{t,x,\alpha} - x|^2 \;\; \geq \;\; r^2 \;\; > \;\; 0 \ .$$

Therefore, we get

$$v(t,x) \;\; \geq \;\; r^2 + \mathbb{E}\Big[v\big(\theta, X_\theta^{t,x,\alpha}\big) + \int_t^\theta f(X_s^{t,x,\alpha}, \alpha_s)ds\Big] \ ,$$

for any control $\alpha \in \mathcal{A}_t$. This contradicts the DPP.