Chapter 7

Optimal control of diffusion processes and dynamic programming

7.1 Stochastic control problem

To define the control problem, we introduce the following elements.

Set of controls We fix a subset A of some $\mathbb{R}^p$ for $p \geq 1$. We suppose that A is bounded and we denote by $\mathcal{A}$ the set of $\mathbb{F}$-progressive processes $(\alpha_t)_{t\in[0,T}$ valued in A.

Controlled diffusion process We fix two functions $b, \sigma : \mathbb{R}^n \times A \to \mathbb{R}^n, \mathbb{R}^{d\times n}$. We suppose that b and σ are continuous and there exists a constant L such that

$$|b(x,a) - b(x',a')| + |\sigma(x,a) - \sigma(x',a')| \quad \leq \quad L\big(|x-x'| + |a-a'|\big) \ ,$$

for all $(x,a),(x',a') \in \mathbb{R}^n \times A$. From Theorem 6.1.1, we have existence and uniqueness of the controlled process $X^{t,x,\alpha}$ defined by

$$X_s^{t,x,\alpha} \quad = \quad x + \int_t^s b(X_u^{t,x,\alpha}, \alpha_u)du + \int_t^s \sigma(X_u^{t,x,\alpha}, \alpha_u)dW_u \qquad (7.1.1)$$

for any $(t,x) \in [0,T] \times \mathbb{R}^n$ and any control process $\alpha \in \mathcal{A}$. Following the same arguments as in the proof of Theorem 6.1.2, we get the following result.

Proposition 7.1.3 *For $p \geq 1$, there exists a constant C_p such that*

$$\begin{aligned}
\mathbb{E}\Big[\sup_{s\in[t,T]}|X_s^{t,x,\alpha}|^p\Big] &\leq C_p(1+|x|^p)\,,\\
\mathbb{E}\Big[\sup_{s\leq t\leq T}|X_s^{t,x,\alpha}-X_s^{t,x',\alpha}|^p\Big] &\leq C_p|x-x'|^p\,,\\
\mathbb{E}\Big[\sup_{t\leq s\leq T}|X_s^{t,x,\alpha}-X_{s\vee(t+h)}^{t+h,x,\alpha}|^p\Big] &\leq C_p h^{\frac{p}{2}}\big(1+|x|^p\big)\,,\\
\mathbb{E}\Big[\sup_{s\leq t\leq T}|X_s^{t,x,\alpha}-X_s^{t,x,\alpha'}|^p\Big] &\leq C_p\mathbb{E}\Big[\int_t^T|\alpha_s-\alpha'_s|^p ds\Big]\,,
\end{aligned}$$

for all $t \in [0,T]$, $h \in [0,T-t]$, $x, x' \in \mathbb{R}^n$ and $\alpha, \alpha' \in \mathcal{A}$.

Reward functions and gain We fix two reward functions $f:\ \mathbb{R}^n \times A \to \mathbb{R}$ and $g:\ \mathbb{R}^n \to \mathbb{R}$. We assume that f and g are locally Lipschitz continuous, that is, for any $N > 0$, there exists a constant L_N such that

$$|f(x,a)-f(x',a')|+|g(x)-g(x')| \;\leq\; L_N\big(|x-x'|+|a-a'|\big)$$

for any $(x,a),(x',a') \in \mathbb{R}^n \times A$ such that $|x| \leq N$ and $|x'| \leq N$. We also assume that f and g have a polynomial growth, that is, there exist a constant C and an integer p such that

$$|f(x,a)|+|g(x)| \;\leq\; C\big(1+|x|^p\big)$$

for all $(x,a) \in \mathbb{R}^n \times A$. We next define the gain functional $J:\ [0,T] \times \mathbb{R}^n \times \mathcal{A} \to \mathbb{R}$ by

$$J(t,x,\alpha) \;=\; \mathbb{E}\Big[g(X_T^{t,x,\alpha})+\int_t^T f(X_s^{t,x,\alpha},\alpha_s)ds\Big]$$

for $(t,x,\alpha) \in [0,T] \times \mathbb{R}^n \times \mathcal{A}$. We observe that $J(t,x,\alpha)$ is well defined under the polynomial growth assumption from Proposition 7.1.3.

We now define the value function v of the considered stochastic control problem by

$$v(t,x) \;=\; \sup_{\alpha\in\mathcal{A}_t} J(t,x,\alpha) \tag{7.1.2}$$

for $(t,x) \in [0,T] \times \mathbb{R}^n$ where

$$\mathcal{A}_t \quad = \quad \{\alpha \in \mathcal{A} \ : \ \alpha \text{ independent of } \mathcal{F}_t\} \ , \quad t \in [0,T] \ .$$

Our goal is to characterize the function v in terms of a partial differential equation called the Hamilton-Jacobi-Bellman (HJB) equation.

Remark 7.1.1 *It can be proved that v satisfies*

$$v(t,x) \quad = \quad \sup_{\alpha \in \mathcal{A}} J(t,x,\alpha)$$

for $(t,x) \in [0,T] \times \mathbb{R}^n$. Indeed, as the reward $J(t,x,.)$ involves only randomness independent of $\mathcal{F}_t$, one can show by a conditioning property that a control in $\mathcal{A}$ provides the same reward as a control in $\mathcal{A}_t$. This property is admitted.

7.2 Dynamic programming principle

To state the dynamic programming principle, we first provide a regularity result on the function v.

Proposition 7.2.4 *For a compact set $\Theta \subset [0,T] \times \mathbb{R}^n$ there exists a real map $\lambda_\Theta : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\lambda_\Theta(r) \to 0$ as $r \to 0$ and*

$$|J(t,x,\alpha) - J(s,y,\alpha)| \quad \leq \quad \lambda_\Theta\big(|t-s| + |x-y|\big) \tag{7.2.3}$$

for all $(t,x),(s,y) \in \Theta$ and $\alpha \in \mathcal{A}$. The value function v is locally uniformly continuous and has polynomial growth.

Proof. Since f and g are locally Lipschitz continuous with polynomial growth, we get from Proposition 7.1.3 that there exist a constant C, an integer p such that for each N, there exists a constant C_N such that

$$\begin{aligned} & |J(t,x,\alpha) - J(s,y,\alpha)| \\ \leq \quad & C_N|t-s|^{\frac{1}{2}}(1+|x|) + L_N\big(|x-y| + |s-t|\big) \\ & + C(1+|x|^p + |y|^p)\mathbb{P}\Big(\sup_{u \in [t,T]} \min\{|X_u^{t,x,\alpha}|, |X_{u \vee s}^{s,y,\alpha}|\} \geq N\Big) \ , \end{aligned}$$

for any $t, s \in [0,T]$ $t \leq s$, any $x, y \in \mathbb{R}^n$ such that $(t,x),(s,y) \in \Theta$ and any $\alpha \in \mathcal{A}$. Using Markov inequality, we get

$$\begin{aligned} & \sup_{\alpha\in\mathcal{A}} |J(t,x,\alpha) - J(s,y,\alpha)| \\ \leq \ & C_N|t-s|^{\frac{1}{2}}(1+|x|) + L_N\big(|x-y|+|s-t|\big) \\ & +C(1+|x|^p+|y|^p)/N \ . \end{aligned}$$

which gives 7.2.3 and the local uniform continuity of v from Remark 7.1.1. The polynomial growth of v follows from the polynomial growth of f and g and Proposition 7.1.3. □

We can now state the DPP.

Theorem 7.2.4 (Dynamic programming principle) *The value function v satisifies*

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

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 $\alpha^\omega \in \mathcal{A}_{\theta(\omega)}$ such that

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

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

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

Step 2. We now prove the reverse inequality. Fix $(t_0, x_0) \in [0,T) \times \mathbb{R}^n$, $\alpha \in \mathcal{A}_{t_0}$, 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 $\alpha^n \in \mathcal{A}_{t_n}$ such that

$$J(t_n, x_n, \alpha^n) \quad \geq \quad v(t_n, x_n) - \varepsilon \tag{7.2.4}$$

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

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

for some $\eta > 0$ and

$$|v(.) - v(t_n, x_n)| + |J(., \alpha_n) - J(t_n, x_n, \alpha_n)| \quad \leq \quad \varepsilon \text{ on } B_n\ . \tag{7.2.5}$$

Let us now define the stopping time

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

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

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

Since α^n is independent of $\mathcal{F}_{t_n}$ for each $n \geq 1$, we get from (7.2.4) and (7.2.5)

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

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

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

□

7.3 Dynamic programming 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

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

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

$$\sup_{a\in A}\big\{\mathcal{L}^a v(t,x) + f(x,a)\big\} \;\; = \;\; 0 \, , \quad (t,x) \in [0,T) \times \mathbb{R}^n \qquad (7.3.6)$$

together with the terminal condition

$$v(T,x) \;\; = \;\; g(x) \, , \quad x \in \mathbb{R}^n \, . \qquad (7.3.7)$$

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

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

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

Theorem 7.3.5 *Suppose that $v \in C^0([0,T] \times \mathbb{R}^n) \cap C^{1,2}([0,T) \times \mathbb{R}^n)$. Then, v is a solution to* (7.3.6)-(7.3.7).

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) and $\eta > 0$ such that

$$\mathcal{L}^a\varphi(.) + f(.,a) \quad > \quad \eta \text{ on } V \; . \tag{7.3.8}$$

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

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

In view of (7.3.8) 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

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

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

$$\mathcal{H}w \quad \leq \quad 0 \text{ on } V \tag{7.3.9}$$

where the function w is defined by

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

For $\alpha \in \mathcal{A}_t$, let θ be the first exit time of $(s, X_s^{t,x,\alpha})$ from V. Using Itô's formula and (7.3.9), we have

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

From the definition of w and θ, it follows that

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

We then notice from the definition of θ that

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

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. $\square$