Chapitre 4 : Équation de Hamilton-Jacobi-Bellman

Introduction : Ce chapitre est consacré à la programmation dynamique en temps continu, nous allons introduire une fonction valeur $V : [t_0, T] \times \mathbb{R}^n \to \mathbb{R}$ et nous allons montrer grâce, au principe d'optimalité de Bellman que cette fonction est solution d'une équation aux dérivées partielles en espace et en temps, appelée équation de Hamilton-Jacobi-Bellman (HJB).

Nous allons également montrer, sous certaines hypothèses, comment la fonction valeur nous permet de synthétiser un contrôle sous forme de feedback

$$U_{adm} = L^1([t_0, T]; U)$$

$$= \left\{ u : [t_0, T] \to U \,/\, \int_{t_0}^{T} |u(t)|\, dt < +\infty \right\}.$$

On considère une fonction $g : [t_0, T] \times \mathbb{R}^n \times U \to \mathbb{R}$ et une fonction $\Psi : \mathbb{R}^n \to \mathbb{R}$ et on définit le critère suivant:

$$J(t_0, x_0, u) = \int_{t_0}^{T} g(t, x, u)\, dt + \Psi(x(T)).$$

Le problème de contrôle optimal que l'on considère est le suivant:. Chercher $u^* \in U_{adm}$ tel que

$$J(t_0, x_0, u^*) = \inf_{u \in U_{adm}} J(t_0, x_0, u) \qquad (4.2)$$

Maintenant on plonge le problème de minimisation (4.2) ~~dans~~ dans une famille de problèmes de contrôle

$\boxed{4.3°}$

$$U_{adm} = L^1([t_0, T]; U)$$

$$= \left\{ u : [t_0, T] \to U \,/\, \int_{t_0}^{T} |u(t)|\,dt < +\infty \right\}.$$

On considère une fonction $g : [t_0, T] \times \mathbb{R}^n \times U \to \mathbb{R}$ et une fonction $\Psi : \mathbb{R}^n \to \mathbb{R}$ et on définit le critère suivant :

$$J(t_0, x_0, u) = \int_{t_0}^{T} g(t, x, u)\,dt + \Psi(x(T)).$$

Le problème de contrôle optimal que l'on considère est le suivant : chercher $u_* \in U_{adm}$ tel que

$$J(t_0, x_0, u_*) = \inf_{u \in U_{adm}} J(t_0, x_0, u) \qquad (4.2)$$

Maintenant on plonge le problème de minimisation (4.2) dans une famille de problèmes de contrôle optimal paramétrés par la paire (S, ξ) où $S \in [t_0, T]$ et $\xi \in \mathbb{R}^n$

Ces paramètres nous indiquent que le problème de contrôle optimal paramétré par (S, ξ) est posé sur l'intervalle $I_S = [S, T]$ avec la condition initiale $X(S) = \xi$.

En résumé, on considère donc la famille de systèmes de contrôle non-linéaires et de critères

$$\left\{ \begin{array}{l} \dot{x}(t) = f(t, x(t), u(t)), \ t \in [S, T], \\ X(S) = \xi, \\ J(S, \xi; u) = \int_S^T g(t, x, u)\, dt + \Psi(x(T)), \ \forall u \in \tilde{U}_{adm} \end{array} \right. \quad (4.3)$$

où $\tilde{U}_{adm} = L^1(I_S, U)$ et $I_S = [S, T]$.

<u>Définition 4.1</u> : Fonction valeur

La fonction valeur $V : [t_0, T] \times \mathbb{R}^n \to \mathbb{R}$ associée à la famille de système de contrôle non-linéaires et de critères (4.3) est telle que

$\boxed{4.4^0}$

$$V(s,\xi) = \inf_{u \in \mathcal{U}_s} J(s,\xi;u). \quad (4.4)$$

En $s = T$, on pose $V(T,\xi) = \Psi(\xi)$ pour tout $\xi \in \mathbb{R}^n$.

Dans la suite de ce chapitre, nous faisons les hypothèses suivantes sur la fonction f

H1) La fonction f est continue sur $[t_0, T] \times \mathbb{R}^n \times U$, uniformément en $u \in U$;

H2) La fonction f est dérivable par rapport à x et la fonction $\dfrac{\partial f}{\partial x}$ est continue bornée sur $[t_0, T] \times \mathbb{R}^n \times U$;

H3) il existe une constante $c \geq 0$ telle que

$$\| \underbrace{f(t,x,\cancel{B})}_{f(t,x,u)} \|_{\mathbb{R}^n} \leq C\left(1 + \|x\|_{\mathbb{R}^n} + \|u\|_{\mathbb{R}^\ell}\right)$$

sur $[t_0, T] \times \mathbb{R}^n \times U$.

$\boxed{4.5^\circ}$

et les hypothèses suivantes sur le critère

H4) la fonction g est continue sur $[t_0, T] \times \mathbb{R}^n \times U$, uniformément en $u \in U$;

H5) il existe des constantes $\nu > 0$ et $\mu \geq 0$ telles que

$$g(t, x, u) \geq \nu \|u\|^2_{\mathbb{R}^k} - \mu \text{ sur } [t_0, T] \times \mathbb{R}^n \times U;$$

H6) la fonction Ψ est continue et minorée sur $\mathbb{R}^n$.

On pourra vérifier que les hypothèses ci-dessus sont des conditions suffisantes d'une part qu'il existe une unique trajectoire $x \in AC([t_0, T]; \mathbb{R}^n)$ pour tout contrôle $u \in \mathcal{U}_{adm} = L^1([t_0, T]; U)$, et d'autre part pour que le critère J ait bien un sens et $V(s, \xi) > -\infty$ pour tout $(s, \xi) \in [t_0, T] \times \mathbb{R}^n$.

<u>Remarque</u> : Non différentiabilité de la fonction valeur

On fera attention au fait que la fonction valeur n'est pas toujours différentiable en tout point.

Passons maintenant au principe d'optimalité de Bellman. Celui-ci constitue la clé de voute de la programmation dynamique. Heuristiquement, le principe d'optimalité de Bellman nous dit que un contrôle u^* est optimal si à tout instant $s \in [t_0, T]$ sur la trajectoire associée x^* le contrôle restreint aux instants ultérieurs $u^*|_{[s,T]}$ est optimal pour le nouveau problème ayant l'état courant $x^*(s)$ comme état initial.

<u>**Théorème P.P.B**</u> (Principe d'optimalité de Bellman).

Soit $u^* \in U_{adm}$ un contrôle optimal. Soit $(s, \xi) \in [t_0, T] \times \mathbb{R}^n$. Alors pour tout $s' \in [s, T]$, on a

$$V(s, \xi) = \inf_{u \in U_s^{s'}} \left\{ \int_{I_s^{s'}} g(t, x(t), u(t))\, dt + V(s', x(s')) \right\}, \quad (4.5)$$

où ~~avec~~ $U_s^{s'} = L^1(I_s^{s'}, U)$ avec $I_s^{s'} = [s, s']$.

<u>Preuve</u> : On observe que toute fonction $u \in \mathcal{U}_S$ peut être identifiée à un couple de fonctions $(u_1, u_2) \in \mathcal{U}_S^{S'} \times \mathcal{U}_{S'}$ en posant $u_1 = u\Big|_{I_S^{S'} = [S, S']}$ et $u_2 = u\Big|_{I_{S'} = [S', T]}$.

D'une part le contrôle u_1 conduit à la trajectoire $x_1 \equiv x$ sur $I_S^{S'}$ telle que $x_1(S) = \xi$.

D'autre part le contrôle u_2 conduit à la trajectoire $x_2 \equiv x$ sur $I_{S'}$ telle que $x_2(S') = x_1(S')$.

On a,

$$V(S, \xi) \overset{\text{déf}}{=} \inf_{u \in \mathcal{U}_S} \left(\int_{I_S} g(t, x(t), u(t))\, dt + \Psi(x(T)) \right)$$

$$= \inf_{(u_1, u_2) \in \mathcal{U}_S^{S'} \times \mathcal{U}_{S'}} \left(\int_{I_S^{S'}} g(t, x_1(t), u_1(t))\, dt + \int_{I_{S'}} g(t, x_2(t), u_2(t))\, dt + \Psi(x_2(T)) \right)$$

$$= \inf_{u_1 \in \mathcal{U}_S^{S'}} \left(\int_{I_S^{S'}} g(t, x_1(t), u_1(t))\,dt + \inf_{u_2 \in \mathcal{U}_{S'}} \left(\int_{I_{S'}} g(t, x_2(t), u_2(t))\,dt + \Psi(x_2(T)) \right) \right)$$

$$= \inf_{u_1 \in \mathcal{U}_S^{S'}} \left(\int_{I_S^{S'}} g(t, x_1(t), u_1(t))\,dt + V(S', x_1(S')) \right),$$

ce qui achève la preuve.

On rappelle que le Hamiltonien associé au système de contrôle non-linéaire (4.1) et au critère (4.2) est l'application H définie par

$$H(t, x, \lambda, u) = \langle \lambda, f(t, x, u) \rangle + g(t, x, u),$$

où $(t, x, \lambda, u) \in [t_0, T] \times \mathbb{R}^n \times \mathbb{R}^n \times U$.

En outre, on définit le Hamiltonien minimisé comme l'application $H_{min} : [t_0, T] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ telle que

$$H_{min}(t, x, \lambda) = \inf_{u \in U} H(t, x, \lambda, u)$$

On a le résultat suivant :

<u>Théorème</u> : Équation HJB (Équation de Hamilton-Jacobi-Bellman).

En tout point $(s, \xi) \in [t_0, T] \times \mathbb{R}^n$ où la fonction valeur est différentiable, elle satisfait l'équation

$$\frac{\partial V}{\partial s}(s, \xi) + H_{min}\left(s, \xi, \frac{\partial V}{\partial \xi}(s, \xi)\right) = 0 \quad (4.6)$$

et elle vérifie en outre la condition à l'instant final

$$V(T, \xi) = \Psi(\xi), \ \forall \xi \in \mathbb{R}^n \quad (4.7)$$

Exemple : On considère le problème de contrôle optimal suivant

$$\text{(Ex.1)}\quad \begin{cases} \dot{x}(t) = u(t) \in [-1,1], \quad t \in [S,T] \\ x(S) = \xi, \\ \inf\limits_{u \in U = [-1,1]} \Psi(x(T)), \end{cases}$$

où $\Psi : \mathbb{R} \to \mathbb{R}$ une fonction.

Le Hamiltonien H associé au problème (Ex.1) est :

$$H(t,x,\lambda,u) = \lambda \cdot u.$$

Alors,

$$H_{min}(t,x,\lambda) = \inf_{u \in [-1,1]} H(t,x,\lambda,u)$$

$$= \inf_{u \in [-1,1]} \lambda \cdot u = \begin{cases} +1 \text{ si } \lambda(t) < 0, \\ -1 \text{ si } \lambda(t) > 0 \end{cases}$$

$$= -|\lambda(t)|.$$

Par suite l'équation HJB est :

$$\frac{\partial V}{\partial s}(s,\xi) + H_{min}\left(s,\xi, \frac{\partial V}{\partial \xi}(s,\xi)\right) = 0,$$

c'est-à-dire,

$$\frac{\partial V}{\partial s}(s,\xi) - \left|\frac{\partial V}{\partial \xi}(s,\xi)\right| = 0.$$

De plus on a la condition à l'instant final

$$V(T,\xi) = \Psi(\xi), \ \forall \xi \in \mathbb{R}.$$

<u>Remarque</u> : Unicité de la solution régulière.

On peut montrer que si une fonction W suffisament régulière (à savoir $W \in C([t_0,T] \times \mathbb{R}^n) \cap C^1([t_0,T[\times \mathbb{R}^n)$) satisfait l'équation HJB (4.6) et la condition en temps final (4.7) et si le sous-ensemble U est borné, alors $W \equiv V$. En d'autres termes, l'équation

(4.12)

HJB (4.6) avec la condition (4.7) a au plus une solution régulière. Ce résultat d'unicité s'étend au cas où le sous-ensemble U est non borné sous hypothèse de décroissance de W quand $\|\xi\| \to +\infty$ uniformément en t.

<u>Proposition</u> : (Synthèse d'un feedback optimal)

On suppose que la fonction valeur V est suffisament régulière, c'est-à-dire, $V \in C([t_0, T] \times \mathbb{R}^n) \cap C^1([t_0, T[\times \mathbb{R}^n)$.

On suppose que pour tout $(s, \xi) \in [t_0, T] \times \mathbb{R}^n$, on peut trouver un feedback optimal

$$u^*(s, \xi) = \min_{v \in U} H(s, \xi, \frac{\partial V}{\partial \xi}(s, \xi), v)$$

(L'existence d'un tel feedback optimal est assuré par les hypothèses sur f et g, en général, on n'a pas unicité, ni dépendance continue en (s, ξ)).

On suppose enfin que l'on peut choisir le feedback $u^*(s,\xi)$ de sorte à ce que le système differentiel

$$\begin{cases} \dot{x}^*(t) = f(t, x^*(t), u^*(t, x^*(t))), \; t \in [t_0, T], \\ x^*(t_0) = x_0, \end{cases}$$

admet une solution $x^* \in AC([t_0, T]; \mathbb{R}^n)$.

Dans ces conditions

$$u^*(t) = u^*(t, x^*(t))$$

est un contrôle optimal sur $[t_0, T]$.

<u>Preuve</u> : On a,

$$\frac{dV}{dt}(t, x^*(t)) = \frac{\partial V}{\partial s}(t, x^*(t)) + \left\langle \frac{\partial V}{\partial \xi}(t, x^*(t)), f(t, x^*(t), u^*(t)) \right\rangle, \quad \text{p.p. sur } [t_0, T]$$

Comme la fonction V satisfait l'équation HJB, on a

$$\frac{\partial V}{\partial s}(t, x^*(t)) + H_{min}\left(t, x^*(t), \frac{\partial V}{\partial \xi}(t, x^*(t))\right) = 0.$$

Par définition de u^*, on a

$$H_{min}\left(t, x^*(t), \frac{\partial V}{\partial \xi}(t, x^*(t))\right)$$

$$= H\left(t, x^*(t), \frac{\partial V}{\partial \xi}(t, x^*(t)), u^*(t)\right)$$

Comme
$$H(t, x, \lambda, u) = \langle \lambda, f(t, x, u) \rangle + g(t, x, u),$$

on obtient

$$\frac{d}{dt} V(t, x^*(t)) = -g(t, x^*(t), u^*(t)) \text{ p.p. sur } [t_0, T].$$

Par suite on déduit que,

$$V(t_0, x_0) = V(t_0, x^*(t_0))$$
$$= V(T, x^*(T)) - \int_{t_0}^{T} \frac{d}{dt} V(t, x^*(t))$$

$$= \Psi(x^*(T)) + \int_{t_0}^{T} g(t, x^*(t), u^*(t))\, dt$$

$$= J(t_0, x_0, u^*),$$

ce qui montre que u^* est bien un contrôle optimal.

<u>Proposition</u> : Fonction valeur et état adjoint

On suppose qu'il existe un contrôle optimal $u^* : [t_0, T] \to U$.

On note $x^* : [t_0, T] \to \mathbb{R}^n$ la trajectoire correspondante.

Soit $\lambda : [t_0, T] \to \mathbb{R}^n$ l'état adjoint dans le PMP

(Principe du maximum de Pontryaguine) c'est-à-dire

tel que

$$\dot{\lambda}(t) = -\frac{\partial H}{\partial x}(t, x^*(t), \lambda(t), u^*(t))$$

$$= -\left\langle \frac{\partial f}{\partial x}(t, x^*(t), u^*(t)), \lambda(t) \right\rangle - \frac{\partial g}{\partial x}(t, x^*(t), u^*(t)),$$

pour tout $t \in [t_0, T]$,

et $$\lambda(T) = \frac{\partial \Psi}{\partial x}(x^*(T)).$$

On suppose que la fonction valeur V est différentiable en $(s, x^*(s))$ pour tout $s \in [t_0, T]$. Dans ces conditions, on a

$$\lambda(s) = \frac{\partial V}{\partial \xi}(s, x^*(s)), \text{ pour tout } s \in [t_0, T].$$

Exemples d'applications

<u>Exemple 1</u> : Mouvement d'un point matériel.

On considère le mouvement d'un point matériel avec un critère quadratique

$$\begin{cases} \dot{x}(t) = u(t), \quad t \in [s, T], \\ x(s) = \xi, \\ J(s, \xi; u) = \frac{1}{2}\int_s^T \left(u^2(t) + x^2(t)\right) dt, \end{cases}$$

pour tout $s \in [0, T]$, $T > 0$ fixé, et pour tout $\xi \in \mathbb{R}$.

<u>Question 1</u> : Écrire l'équation HJB et la condition finale pour la fonction valeur $V(s, \xi) = \inf_{u \in L^1([0,T];\mathbb{R})} J(s, \xi; u)$.

Réponse :

Le Hamiltonien H est :

$$H(x,\lambda,u) = \lambda . u + \frac{1}{2}(x^2+u^2).$$

Alors le Hamiltonien minimisé H_{min} est

$$\begin{aligned} H_{min}(x,\lambda) &= \inf_{u\in\mathbb{R}} H(x,\lambda,u) \\ &= \inf_{u\in\mathbb{R}} \left\{\lambda . u + \frac{1}{2}(x^2+u^2)\right\}. \\ &= H(x,\lambda,-\lambda) \text{ avec } u=-\lambda \text{ comme unique minimiseur.} \\ &= \lambda.(-\lambda)+\frac{1}{2}\left(x^2+(-\lambda)^2\right) \\ &= \frac{1}{2}(x^2-\lambda^2). \end{aligned}$$

Alors l'équation HJB est :

$$\begin{cases} \frac{\partial V}{\partial s}(s,\xi) + H_{min}\left(\xi, \frac{\partial V}{\partial \xi}(s,\xi)\right) = 0, \\ V(T,\xi) = 0 \quad \text{car } \Psi(\xi)=0. \end{cases}$$

$\boxed{4.18°}$

C'est-à-dire,

$$\begin{cases} \dfrac{\partial V}{\partial s}(s,\xi) + \dfrac{1}{2}\left(\xi^2 - \left(\dfrac{\partial V}{\partial \xi}(s,\xi)\right)^2\right) = 0, \\ V(T,\xi) = 0. \end{cases}$$

<u>Question 2</u>: Résoudre l'équation HJB en cherchant la solution sous la forme séparée $V(s;\xi) = \frac{1}{2}\mu(s)\xi^2$.

<u>Réponse</u> : On a,

$$V(s;\xi) = \frac{1}{2}\mu(s)\xi^2$$

Alors,
$$\frac{\partial V}{\partial s}(s,\xi) = \frac{1}{2}\mu'(s)\xi^2,$$

et
$$\frac{\partial V}{\partial \xi}(s,\xi) = \mu(s)\xi.$$

Par suite d'après l'équation HJB, on obtient

$$\begin{cases} \frac{1}{2}\mu'(s)\xi^2 + \frac{1}{2}\left(\xi^2 - (\mu(s)\xi)^2\right) = 0, \\ \mu(T) = 0. \end{cases}$$

C'est-à-dire,

$$\begin{cases} \mu'(S) + 1 - \mu^2(S) = 0, \\ \mu(T) = 0. \end{cases}$$

On a,

$$\mu'(S) = \mu^2(S) - 1.$$

Alors,

$$\int \frac{d\mu}{\mu^2 - 1} = \int dS \quad (\text{voir que } \mu \neq \pm 1 \text{ car } \mu(T) = 0).$$

C'est-à-dire,

$$\frac{1}{2} \int \left(\frac{1}{\mu - 1} - \frac{1}{\mu + 1} \right) d\mu = \int dS.$$

C'est-à-dire,

$$\frac{1}{2} \ln \left| \frac{\mu - 1}{\mu + 1} \right| = S + k, \quad \text{avec } k \in \mathbb{R}$$

Ce qui donne,

$$\frac{\mu(S) - 1}{\mu(S) + 1} = K e^{2S}, \quad \text{avec } K \in \mathbb{R}^*.$$

$\boxed{4.20°}$

C'est-à-dire,

$$\mu(s) - 1 = Ke^{2s}(\mu(s) + 1)$$

Ce qui donne,

$$\mu(s) = \frac{Ke^{2s} + 1}{1 - Ke^{2s}}$$

Maintenant comme $\mu(T) = 0$, on obtient

$$\frac{Ke^{2T} + 1}{1 - Ke^{2T}} = 0$$

c'st-à-dire, $Ke^{2T} + 1 = 0$

C'st-à-dire, $K = -e^{-2T}$.

Par suite,

$$\mu(s) = \frac{1 - e^{2(s-T)}}{1 + e^{2(s-T)}}$$

$$= th(T - s).$$

$\boxed{4.21°}$

Alors,

$$V(s,\xi) \overset{\text{déf}}{=} \frac{1}{2}\mu(s)\xi^2$$

$$= \boxed{\frac{\operatorname{th}(T-s)}{2}\xi^2}$$

<u>Question 3</u> : En déduire le contrôle optimal comme feedback, la trajectoire optimale et le contrôle optimal.

<u>Réponse</u> :

i) Le contrôle optimal sous forme de feedback est

$$\tilde{u}(s,\xi) = -\frac{\partial V}{\partial \xi}(s,\xi) = -\operatorname{th}(T-s)\xi.$$

ii) La trajectoire optimale x^* :

x^* est solution de

$$\left\{\begin{array}{l} \dot{x}^*(t) = -\operatorname{th}(T-t)\, x^*(t), \\ x^*(t_0) = x_0. \end{array}\right. \qquad \Sigma$$

Alors

$$\ln|x^*(t)| - \ln|x_0| = -\int_0^t \operatorname{th}(T-s)\,ds \quad \text{si } x_0 \neq 0$$

$$= -\int_0^t \frac{\operatorname{sh}(T-s)}{\operatorname{ch}(T-s)}\,ds.$$

$$= \ln \operatorname{ch}(T-s)\Big|_0^t$$

$$= \ln \operatorname{ch}(T-t) - \ln \operatorname{ch} T.$$

C'st-à-dire,

$$\ln\left|\frac{x^*(t)}{x_0}\right| = \ln \frac{\operatorname{ch}(T-t)}{\operatorname{ch} T}.$$

C'st-à-dire,

$$x^*(t) = \frac{x_0}{\operatorname{ch} T}\operatorname{ch}(T-t) \quad \left(\begin{array}{c}\text{Voir que} \\ x^*(0) = x_0\end{array}\right).$$

D'autre par si $x_0 = 0$, alors $x^* \equiv 0$ est ~~par suite~~ solution de Σ.

En conclusion la solution de Σ est donnée par

$$X^*(t) = \frac{x_0}{ch T} ch(T-t).$$

ii) Le contrôle optimal :

On a, $\quad \dot{X}^*(t) = u^*(t).$

Alors, $\quad u^*(t) = \frac{d}{dt}\left(\frac{x_0}{ch T} ch(T-t)\right)$

$$= \boxed{-\frac{x_0}{ch T} sh(T-t)}$$

<u>Exemple 2</u> : Système linéaire quadratique.

On considère la famille de systèmes linéaires quadratique

$$\begin{cases} \dot{X}(t) = A(t)X(t) + B(t)u(t), \ t \in [s,T] \\ X(s) = \xi, \\ J(s,\xi;u) = {}^tX(T) Q X(T) + \int_0^T \left({}^tX(t) W(t) X(t) + {}^tu(t) U(t) u(t) \right) dt \end{cases}$$

$\boxed{4.24°}$

où $T>0$ est fixé, et où, pour tout $t\in[0,T]$, $U(t)\in M_{k}(\mathbb{R})$ est symétrique définie positive, $W(t)\in M_n(\mathbb{R})$ est symétrique positive, et $Q\in M_n(\mathbb{R})$ est une matrice symétrique positive. On suppose que les dépendances en t de A, B, W et U sont L^∞ sur $[0,T]$. Comme le coût est quadratique, l'espace naturel des contrôles est $L^2([0,T],\mathbb{R}^k)$.

<u>Question 1</u> : Écrire l'équation HJB et la condition finale pour la fonction valeur $V(s,\xi)=\inf\limits_{u\in L^2([0,T],\mathbb{R}^k)} J(s,\xi;u)$.

<u>Réponse</u> : Le Hamiltonien associé au problème LQ (linéaire quadratique) est

$$H(x,\lambda,u)=\langle\lambda, A(t)x+B(t)u\rangle + {}^t x(t)W(t)x(t)+{}^t u(t)U(t)u(t)$$

$$= {}^t\lambda(t)\big(A(t)x(t)+B(t)u(t)\big)+{}^t x(t)W(t)x(t) + {}^t u(t)U(t)u(t).$$

$\boxed{4.25°}$

Alors le Hamiltonien minimisé H_{min} est

$$
\begin{aligned}
& H_{min}(X, \lambda) \\
&= \min_{u \in \mathbb{R}^k} H(X, \lambda, u) \\
&= {}^t\lambda(t) A(t) X(t) - \frac{1}{2}\, {}^t\lambda(t) B(t) U^{-1}(t)\, {}^tB(t) \lambda(t) \\
&\quad + {}^tX(t) W(t) X(t) + \frac{1}{4}\, {}^t\lambda(t) B(t) U^{-1}(t)\, {}^tB(t) \lambda(t) \\
&= {}^t\lambda(t) A(t) X(t) - \frac{1}{4}\, {}^t\lambda(t) B(t) U^{-1}(t)\, {}^tB(t) \lambda(t) \\
&\quad + {}^tX(t) W(t) X(t),
\end{aligned}
$$

avec l'unique minimiseur $u(t) = -\frac{1}{2} U^{-1}(t)\, {}^tB(t) \lambda(t)$.

Par suite l'équation HJB est:

$$\frac{\partial V}{\partial s}(s,\xi) + H_{min}\left(\xi, \frac{\partial V}{\partial \xi}(s,\xi)\right) = 0,$$

c'est-à-dire,

$$\frac{\partial V}{\partial s}(s,\xi) + {}^t\!\left(\frac{\partial V}{\partial \xi}(s,\xi)\right) A(s)\xi - \frac{1}{4}\,{}^t\!\left(\frac{\partial V}{\partial \xi}(s,\xi)\right) B(s)U^{-1}(s)\,{}^tB(s)\frac{\partial V}{\partial \xi}(s,\xi)$$

$$+ {}^t\xi\, W(s)\xi = 0,$$

et la condition finale est

$$V(T,\xi) = \Psi(\xi) = {}^tX(T)\,Q\,X(T).$$

Question 2 : On suppose que $V(s,\xi) = \frac{1}{2}\,{}^t\xi\, P(s)\xi$.
Déterminer l'équation différentielle satisfaite par l'équation HJB.

Réponse : On a,

$$\frac{\partial V}{\partial s}(s,\xi) = \frac{1}{2}\,{}^t\xi\, P'(s)\xi,$$

et

$$\frac{\partial V}{\partial \xi}(s,\xi) = P(s)\xi.$$

En remplace dans l'équation HJB, on obtient

$$^t\xi\left[\frac{1}{2}P'(S)+{}^tP(S)A(S)-\frac{1}{4}{}^tP(S)B(S)U^{-1}(S)\,{}^tB(S)P(S)+W(S)\right]\xi=0,$$

Ce qui entraîne que la partie symétrique de la matrice entre parenthèse est nulle c'est-à-dire

$$P'(S)=-P(S)A(S)-{}^tA(S)P(S)-\frac{1}{2}P(S)B(S)U^{-1}(S)\,{}^tB(S)P(S)$$
$$-2W(S),$$

et la condition finale est $P(T)=Q$.

*) <u>Bilan</u> : PMP ou HJB

Pour résumer les principaux résultats que nous avons vus sur le PMP et l'équation HJB.

Nous pouvons conclure avec les commentaires suivants;

Le PMP :

*) fournit une condition nécessaire d'optimalité;

**) fournit le contrôle optimal en boucle ouverte (fonction de temps);

***) repose sur la résolution d'équations différentielles ordinaires;

****) ne s'applique (sauf rares exceptions) qu'aux systèmes déterministes.

En revanche, la programmation dynamique via la résolution de l'équation HJB

*) fournit une condition suffisante d'optimalité;

**) fournit le contrôle optimal en boucle fermée (~~en~~ fonction de l'état);

***) repose sur la résolution d'une équation aux dérivées partielles (ce qui devient rapidement intractable lorsque la dimension n de l'espace des états croit);

****) s'applique aux systèmes déterministes et stochastiques.

$\boxed{4.29^{\circ}}$