Cours de Décision dans l'incertain
Exercices : 21 mai 2021.

Exercice 1 On se donne une suite de fonctions (ou contrôles) $\tilde{u} = (\tilde{u}(n, x_{0:n}), n \geq 0)$ de E^n à valeurs dans un ensemble A ("on choisit la contrôle $\tilde{u}$ en tenant compte de toute la trajectoire avant l'instant n").

$(W_n, n \geq 1)$ une suite i.i.d. à valeurs dans F et x_0 un point de E. Les espaces E, F et A sont supposés finis.

On pose $X_0 = x_0$, puis itérativement on définit $X_{0:n} = (X_0, \dots, X_n)$, $U_n = \tilde{u}(n, X_{0:n})$ puis X_{n+1} par

$$X_{n+1} = f(X_n, U_n, W_{n+1}). \tag{1}$$

A ce système dynamique contrôlé (1) on associe une famille $(P_u(x, y), x, y \in E, u \in A)$ de matrice de transition, en posant

$$P_u(x, y) = \mathbb{P}(f(x, W_1, u) = y)$$

1. Vérifier que pour un u fixé, $P_u(x, y)$ est une matrice de transition d'un chaine de Markov.
2. Montrer que, pour toute fonction v de E dans $\mathbb{R}$, $\sum_{y \in E} P_u(x, y) v(y) = \mathbb{E}[v(f(x, u, W_1)]$.
3. Montrer que $\mathbb{P}(X_{0:n+1} = x_{0:n+1}) = \mathbb{P}(X_{0:n} = x_{0:n}) P_{\tilde{u}(n, x_{0:n})}(x_n, x_{n+1})$.

Exercice 2 Soit $c(n, x, u)$ une fonction de $\mathbb{N} \times E \times A$ dans $\mathbb{R}$ et $K(x)$ un fonction de E dans $\mathbb{R}$. On considère l'équation de (Hamilton-Jacobi-)Bellman suivante

$$\begin{cases} v(n, x) = \min_{u \in A} \left\{ \sum_{y \in E} P_u(x, y) v(n + 1, y) + c(n, x, u) \right\}, \ n < N, \\ v(N, x) = K(x). \end{cases} \tag{HJB}$$

Montrer que l'équation précédente définit une *unique* fonction v. Ecrire un algorithme calculant cette fonction v.

Exercice 3 On spécifie le coût d'une stratégie $\tilde{u}$, en suppose que

— le choix de $U_j = \tilde{u}(j, X_{0:j})$ à l'instant j, coûte $c(j, X_j, U_j)$,

— le coût à l'instant final N est donné par $K(X_N)$.

Le coût moyen d'une stratégie $\tilde{u}$ est alors défini par

$$J(\tilde{u}) = \mathbb{E}\left(\sum_{j=0}^{N-1} c(j, X_j, U_j) + K(X_N) \right).$$

On cherche à identifier une stratégie qui minimise $J(\tilde{u})$, parmi tous les contrôles possibles $\tilde{u}$.

Pour cela on considère $\hat{u}(n, x)$ une fonction (pas forcément unique) à valeurs dans A qui réalise le minimum en u présent dans l'équation (HJB) à n et x fixés. On note $\hat{X}$ le processus définit par réccurence à partir de $\hat{u}$: $X_0 = x_0$, puis

$$\hat{X}_{n+1} = f(\hat{X}_n, \hat{U}_n, W_{n+1}), \text{ où } \hat{U}_n = \hat{u}(n, \hat{X}_n).$$

Le but est de montrer que $\hat{u}$ réalise le minimum de de J.

1. Vérifier que l'on peut définir une fonction $\hat{u}$ qui réalise le minimum en u présent dans l'équation (HJB) mais que cette fonction n'est pas forcément unique.
2. Montrer, en utilisant la question 3 du premier exercice, que

$$\mathbb{E}\left(v(n+1,X_{n+1})\right) = \sum_{x_{0:n}\in E^{n+1}} \mathbb{P}(X_{0:n}=x_{0:n}) \sum_{x_{n+1}\in E} P_{\tilde{u}(n,x_{0:n})}(x_n,x_{n+1})v(n+1,x_{n+1}),$$

puis, en utilisant l'équation (HJB), en déduire que

$$\mathbb{E}(v(n,X_n) - c(n,X_n,U_n)) \leq \mathbb{E}(v(n+1,X_{n+1})). \tag{2}$$

3. En se souvenant que $\hat{u}(n,x)$ réalise le minimum présent dans l'équation (HJB), montrer que

$$\mathbb{E}(v(n,\hat{X}_n) - c(n,\hat{X}_n,\hat{u}(n,\hat{X}_n))) = \mathbb{E}(v(n+1,\hat{X}_{n+1})). \tag{3}$$

4. Montrer par récurrence, en utilisant (2) et $v(N,x)=K(x)$, que :

$$v(0,x_0) \leq \mathbb{E}\left(\sum_{j=0}^{N-1} c(j,X_j,U_j) + K(X_N)\right) = J_{\tilde{u}},$$

5. Montrer à partir de (3) que $v(0,x_0) = \mathbb{E}\left(\sum_{j=0}^{N-1} c(j,\hat{X}_j,\hat{u}(j,\hat{X}_j)) + K(\hat{X}_N)\right) = J_{\hat{u}}$.
6. En déduire l'optimalité de $\hat{u}$: $\min_{\tilde{u}} J(\tilde{u}) = J(\hat{u})$.
   On constate que le contrôle $\hat{u}$ est optimal parmi tous les contrôles mais est de type "feedback" (il ne depend de la trajectoire avant n que par sa dernière valeur X_n).