TD4. Chaînes de Markov contrôlées.

Exercice 1. (Exercer une option d'achat) On a la possibilité d'acheter un actif à un prix fixé d'avance p et à un instant quelconque $n = 0, \ldots, N-1$. Le prix de marché de l'actif est modélisé par une suite $(Y_n)_{n\geqslant 0}$ donnée par $Y_{n+1} = Y_n + \varepsilon_{n+1}$ où $(\varepsilon_n)_{n\geqslant 1}$ est une suite iid intégrable. L'objectif est de maximiser le gain moyen relatif à l'utilisation de l'option d'achat: si on décide de l'utiliser au temps n avec un prix de marché Y_n alors notre gain serait de $Y_n - p$. On modélisera le problème avec un système dynamique contrôlé homogène sur l'espace d'états $\mathcal{M} = \mathbb{R} \cup \{\Delta\}$: à un instant déterminé soit on possède encore l'option et sa valeur est $x \in \mathbb{R}$, soit on a déjà exercé l'option et alors on décide de façon conventionnelle d'être dans l'état fictif Δ. L'espace des actions est $\mathcal{A} = \{0, 1\}$, 0 si on n'exerce pas et 1 si on décide d'exercer l'option.

a) Déterminer la fonction $G: \mathcal{A} \times \mathcal{M} \times \mathbb{R} \to \mathcal{M}$ qui intervient dans la définition de système dynamique contrôlé par rapport à la suite iid $(\varepsilon_n)_{n\geqslant 1}$.

b) Déterminer une fonction de gain homogène $r: \mathcal{M} \times \mathcal{A} \to \mathbb{R}_+$ associé au problème.

c) Montrer que la fonction valeur satisfait

$$V(k,x) = \max\{x - p, \mathbb{E}[V(k+1, x+\varepsilon)]\}, \qquad 0 \leqslant k \leqslant N-1, x \in \mathbb{R}$$

et $V(N,x) = 0$ si $x \in \mathbb{R}$.

d) Par récurrence rétrograde montrer que $V(k, x)$ est une fonction convexe de x pour tout $0 \leqslant k \leqslant N$.

e) Par récurrence rétrograde montrer que $V(k,x) \geqslant V(k+1,x)$ pour tout $0 \leqslant k \leqslant N-1$.

f) Soit $p_k = \inf\{x \geqslant 0: V(k,x) = x - p\}$. Montrez que p_k est décroissant en k.

g) Montrer que la politique optimale est d'exercer l'option de que $Y_k \geqslant p_k$.

Exercice 2. (À la poste) On est en queue à la poste. Avant nous il y a $x \in \mathbb{N}$ personnes et à chaque pas de temps il y a une probabilité $p \in]0, 1[$ que la queue avance d'une position. On estime que l'utilité d'arriver à être servi soit $r > 0$ et que attendre dans la queue nous coût $c > 0$ à chaque instant du temps. On a toujours la possibilité de partir da la queue et de ne pas être servit. On veut trouver une stratégie (a savoir si on doit rester dans la queue et attendre ou partir) qui maximise notre gain moyen.

On modélisera le problème avec un système dynamique contrôlé homogène sur l'espace d'états $\mathcal{M} = \mathbb{N} \cup \{\Delta\}$: à un instant déterminé soit on est dans la queue et on a $x \in \mathbb{N}$ clients avant nous, soit on est parti et alors on décide de façon conventionnelle d'être dans l'état fictif Δ. L'espace des actions est $\mathcal{A} = \{0, 1\}$, 0 si on reste dans la queue pour un autre interval de temps et 1 si on décide de partir.

a) Déterminer la fonction $G: \mathcal{A} \times \mathcal{M} \times \mathbb{R} \to \mathcal{M}$ qui intervient dans la définition de système dynamique contrôlé par rapport à la suite iid $(\varepsilon_n)_{n\geqslant 1}$.

b) Déterminer une fonction de gain homogène $r: \mathcal{M} \times \mathcal{A} \to \mathbb{R}_+$ associé au problème.

c) Déterminer l'équation de Bellman et montrer que on a

$$V(x) = (V(x-1) - c/p)_+, \qquad x \geqslant 1.$$

d) Montrer par récurrence que $V(x) \leqslant V(x-1)$. Donner une explication intuitive de cette inégalité.

e) Montrer qu'il existe $x^* \in \mathbb{N}$ tel que la stratégie optimale est de rester ssi $x \leqslant x^*$.

f) Déterminer $V(x)$ et x^* en fonction de r, c, p.