Feuille d'exercices 4. **Contrôle des chaînes de Markov**

Exercice 1. [Options américaines]

Une option américaine d'achat donne la possibilité (mais pas l'obligation) d'acheter un actif à un prix fixé d'avance p et à un instant quelconque $n = 0, \dots, N-1$. On ne peut utiliser qu'une seule fois une option. Le prix du marché de l'actif est modélisé par un processus $(Y_n)_{n \geqslant 0}$ défini récursivement par

$$Y_{n+1} = Y_n + \varepsilon_{n+1}$$

où $(\varepsilon_n)_{n \geqslant 1}$ est une suite de v.a. i.i.d. intégrables. 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 alors notre gain sera égal à $Y_n - p$.

On modélisera le problème avec un système dynamique contrôlé homogène sur l'espace d'états $S = \mathbb{R} \cup \{\Delta\}$: à un instant déterminé soit on possède encore l'actif 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.

1. Déterminer un espace d'actions A associé au problème.
2. Déterminer la fonction $G : S \times A \times \mathbb{R} \to S$ qui intervient dans la définition du système dynamique contrôlé par rapport à la suite i.i.d. $(\varepsilon_n)_{n \geqslant 1}$.
3. Déterminer une fonction de gain $r : \mathbb{N} \times S \times A \to \mathbb{R}^+$ associée au problème.
4. 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 ε une variable générique de même loi que ε_1 ; et que $V(N,x) = 0$ si $x \in \mathbb{R}$.
5. Par récurrence rétrograde montrer que, pour tout $0 \leq k \leq N$, $x \in \mathbb{R} \mapsto V(k,x)$ est une fonction convexe, et que $x \in \mathbb{R} \mapsto V(k,x) - x$ est une fonction décroissante.
6. Par récurrence rétrograde montrer que, pour tout $x \in \mathbb{R}$, pour tout $0 \leq k \leq N-1$, $V(k,x) \geqslant V(k+1,x)$.
7. Pour tout $0 \leq k \leq N$, on définit $p_k = \inf\{x \geqslant 0 : V(k,x) = x - p\}$ $(\inf\{\emptyset\} = N)$. Montrez que p_k est décroissant en k.
8. Montrer que la politique optimale est d'exercer l'option dès que $Y_k \geqslant p_k$.

Exercice 2. [Examen 2013 – 14]

On dispose au début de l'année n, $n \in \{0, 1, ..., N\}$ d'un capital $X_n \geq 0$. A chaque début d'année $n \leq N-1$, on choisit de dépenser une proportion $U_n \in [0,1]$ du capital X_n et on épargne le reste $(1 - U_n)X_n$ du capital qui, par le biais d'un rendement aléatoire, se fructifie de façon à devenir $R_n(1 - U_n)X_n$ à la fin de l'année n, avec $R_n \geq 0$. On suppose les $(R_n)_{n\geq 0}$ i.i.d. d'espérance finie m. La N-ième année, on dépense tout ce qu'il nous reste. Notre objectif est d'optimiser la somme totale dépensée.

1. Déterminer l'espace d'états S et l'espace d'actions A associés au problème.
2. Déterminer la fonction $G : S \times A \times \mathbb{R}^+ \to S$ et la suite de v.a. i.i.d. $(\varepsilon_n)_{n\geq 0}$ qui interviennent dans la construction du système dynamique contrôlé associé au problème.
3. Déterminer une fonction de gain $r : \mathbb{N} \times S \times A \to \mathbb{R}$ associée au problème.
4. Rappeler la définition de la fonction valeur $V(n, x), n \in \mathbb{N}, x \geq 0$ et écrire l'équation de Bellman vérifiée par cette fonction. Préciser $V(N, x), x \geq 0$.
5. Calculer la fonction $x \in \mathbb{R}^+ \mapsto V(n, x)$ pour tout n. On distinguera trois cas : $m < 1, m = 1$ et $m > 1$.
6. Dans chacun de ces trois cas, donner un contrôle optimal (partant du temps 0).
7. Commenter les résultats obtenus.

Exercice 3. [Contrôle en horizon infini – Examen 2016]

On considère $(X_n)_{n\geq 0}$, un processus de Markov contrôlé **homogène en temps** sur un espace d'états dénombrable S et caractérisé par :

- un espace d'action A,
- une dynamique $P : S \times A \to \text{Proba}(S)$, où $\text{Proba}(S)$ désigne l'ensemble des probabilités sur S.

1. Rappeler la définition du noyau P associé à la dynamique P dans le cas où il agit sur une fonction $f : \mathbb{N} \times S \to \mathbb{R}^+$ puis, dans le cas où il agit sur une fonction $g : S \to \mathbb{R}^+$.

On considère une fonction de **récompense positive** $r : S \times A \to \mathbb{R}^+$ (il s'agit donc d'une fonction de la position du processus et de l'action choisie mais indépendante du temps).

Pour $k \in \mathbb{N}$, on note $\mathcal{C}_k$ l'espace des contrôles partant du temps k. Pour $u \in \mathcal{C}_0$ et $x \in S$, on définit

$$V^u(x) = \mathbb{E}\left(\sum_{i=0}^{+\infty} r(X_i, U_i)\right), \qquad \text{avec pour tout } i \geq 0,\ U_i = u_i(X_0, \cdots, X_i),$$
$$V(x) = \sup_{u \in \mathcal{C}_0} V^u(x).$$

2. Définir la fonction valeur en horizon infini $W : \mathbb{N} \times S \to \mathbb{R}$. Montrer que, pour tout $k \geq 0$ et tout $x \in S$, $W(k, x) = V(x)$.
   Rappeler l'équation de Bellman satisfaite par V.

On définit pour $n \geq 0$, $u \in \mathcal{C}_0$ et $x \in S$

$$V_n^u(x) = \mathbb{E}\left(\sum_{i=0}^{n-1} r(X_i, U_i)\right), \text{ en particulier } V_0^u \equiv 0,$$
$$V_n(x) = \sup_{u \in \mathcal{C}_0} V_n^u(x).$$

3. Montrer que pour tout $u \in \mathcal{C}_0$ et tout $x \in S$, $(V_n^u(x))_{n \geq 0}$ converge en croissant vers $V^u(x)$. En déduire que pour tout $x \in S$, $(V_n(x))_{n \geq 0}$ converge en croissant vers $V(x)$.
4. On considère le problème en horizon fini $N \in \mathbb{N}$. Rappeler la définition de la fonction valeur $W_N : \mathbb{N} \times S \to \mathbb{R}$ associée au problème. Montrer que, pour tout $0 \leq k+1 \leq N$, $W_N(k+1, x) = V_{N-k}(x)$.
5. Montrer que les fonctions V_n, $n \geq 0$, satisfont la relation de récurrence :

$$V_{n+1}(x) = \sup_{a \in A} (r + PV_n)(x, a), \qquad x \in S,\ a \in A.$$

   Indication : utiliser l'équation de Bellman satisfaite par W_N (en choisissant N adapté).
6. Soit $F : S \to \mathbb{R}^+$ une solution positive de l'equation de Bellman. Montrer que pour tout $n \in \mathbb{N}$, $F \geq V_n$.
   En déduire que V est la plus petite solution positive de l'équation de Bellman. Montrer ensuite que si un contrôle u est tel que V^u satisfait l'equation de Bellman alors u est optimal.

Exercice 4. [Examen 2013 – 14]

Considérons le processus contrôlé d'espace d'états $S = \mathbb{N}$, d'ensemble d'actions $A = \{0, 1\}$ et de probabilité de transition homogène définie sur $S \times A$ par

$$P(x, a) = \delta_{a(x + \mathbb{I}_{\{x \geq 1\}})}$$

(il n'y a donc pas d'aléa !). On se donne comme fonction de gain la fonction r définie sur $S \times A$ par

$$r(x, a) = (1-a)\left(1 - \frac{1}{x}\right) \text{ si } x \in \mathbb{N}^* \quad \text{et} \quad r(0,0) = r(0,1) = 0.$$

On note V la fonction valeur associée à ce système (qui est homogène).

1. Calculer $V(0)$ et montrer que V est une solution de l'équation

$$W(x) = \max\left(1 - \frac{1}{x}; W(x+1)\right), \quad x \in \mathbb{N}^*.$$

2. Montrer que W est une solution de l'équation précédente si et seulement si il existe un $\lambda \geq 1$ tel que

$$W(x) = \lambda, \forall x \in \mathbb{N}^*.$$

3. En déduire la valeur de $V(x)$, $\forall x \in \mathbb{N}$.
4. Montrer que pour tout contrôle $u \in \mathcal{C}_0$, $V^u(x) < 1$ pour tout $x \in \mathbb{N}$.
5. En déduire qu'il n'y a pas de contrôle optimal.

Exercice 5. On fait la 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 est $r > 0$ et qu'attendre dans la queue nous coûte $c > 0$ à chaque pas de temps. On a toujours la possibilité de quitter la queue et de ne pas être servi. On veut trouver une stratégie (à savoir si on doit rester dans la queue et attendre ou bien partir) qui maximise notre gain moyen.

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

1. Déterminer la fonction $G : S \times A \times \{0,1\} \to S$ qui intervient dans la définition du système dynamique contrôlé par rapport à une suite i.i.d. $(\varepsilon_n)_{n \geqslant 1}$ que l'on précisera.
2. Déterminer une fonction de gain homogène $r : S \times A \to \mathbb{R}$ associé au problème.
3. Ecrire dans ce contexte l'équation de Bellman et montrer que l'on a
$$V(x) = (V(x-1) - c/p)_+, \qquad x \geqslant 1.$$
4. Montrer que $V(x) \leqslant V(x-1)$, $\forall x \geq 1$. Donner une explication intuitive de cette inégalité.
5. Déterminer $V(x)$ en fonction de r, c, p et x.
6. Montrer qu'il existe $x^* \in \mathbb{N}$ tel qu'une stratégie optimale est de rester ssi $x \leqslant x^*$. Déterminer x^* en fonction de r, c, p.

Exercice 6. [Contrôle d'une marche – Examen 2017]

Soit N un entier strictement positif et $(X_i)_{i=1,\cdots,N}$ une famille de variables aléatoires i.i.d. de loi $P(X = +1) = P(X = -1) = 1/2$. On note $(\mathcal{F}_i)_{i=1,\cdots,N}$ la filtration engendrée par les $(X_i)_{i=1,\cdots,N}$. On considère la marche aléatoire simple définie par

$$\begin{aligned} S_0 &= 0 \\ S_n &= \sum_{i=1}^{n} X_i \qquad \forall n \in \{1, \cdots, N\}. \end{aligned}$$

Dans la première partie, on cherche à contrôler la marche pour minimiser l'espérance du carré de sa distance à l'origine au temps final N. Dans la seconde partie, on cherche à maximiser sur tous les temps d'arrêt la probabilité de s'arrêter au maximum de la marche sur $\{0, \cdots, N\}$. Les deux parties sont complètement indépendantes.

Partie 1. On considère la marche contrôlée $(W_n)_{0 \leq n \leq N}$ en temps discret et à valeurs dans $\mathbb{R}$ définie de la façon suivante. Au temps initial $W_0 = 0$. Pour tout $0 \leq n \leq N-1$, l'incrément de la marche (entre n et $n+1$) est la somme de X_{n+1} et de l'action d'un agent qui peut déplacer la marche à une distance $d \in \mathbb{R}$ quelconque en payant un coût d^2. Pour tout $0 \leq n \leq N-1$, si l'agent choisit l'action $D_n \in \mathbb{R}$, on a donc

$$W_{n+1} = W_n + X_{n+1} + D_n$$

On cherche à minimiser $\mathbb{E}(W_N^2)$ sur l'ensemble des contrôles.

1. Préciser l'espace d'états S, d'actions A, la dynamique (vue comme un système dynamique aléatoire G dont on précisera la suite des innovations) et la fonction de coût c. Écrire le problème d'optimisation que l'on cherche à résoudre.
2. Pour tout $a > 0$ et $b \in \mathbb{R}$, on définit pour $t \in \mathbb{R}$, $f(t) = at^2 + b$. Pour tout $x \in \mathbb{R}$, minimiser sur $\mathbb{R}$
$$u \to \frac{1}{2}\left(f(u+x+1) + f(u+x-1)\right) + u^2$$
et préciser le point où le minimum est atteint.
3. Définir la fonction valeur V du système et écrire l'équation de Bellman associée au problème.
4. Montrer qu'il existe deux suites $(a_k)_{k=0}^N$ et $(b_k)_{k=0}^N$ que l'on définira par récurrence tels que, pour tout $0 \leq k \leq N$ et $w \in \mathbb{R}$, $V(k,w) = a_k w^2 + b_k$. Puis, préciser un contrôle optimal.

Partie 2. On définit pour $1 \leq n \leq N$, $M_n = \max_{k \leq n} S_k$, le maximum de la marche au temps n. On étudie le problème d'arrêt optimal

$\sup\{P(S_T = M_N),\ T \in \mathcal{T}_N\}$ où $\mathcal{T}_N$ désigne l'ensemble des temps d'arrêt bornés par N.

1. Le processus $(1_{\{M_N = S_k\}})_{0 \leq k \leq N}$ est-t-il adapté ?
2. On définit pour $1 \leq k \leq N$, $Y_k = \mathbb{E}[\mathbb{I}_{\{M_N = S_k\}} | \mathcal{F}_k]$. Montrer que pour tout $T \in \mathcal{T}_N$, $\mathbb{E}(Y_T) = \mathbb{P}(S_T = M_N)$.
3. Montrer que pour tout $1 \leq k \leq N$, $Y_k = 1_{\{S_k = M_k\}} p_{N-k}$ avec
$$p_n = \mathbb{P}(S_1 \leq 0, \cdots, S_n \leq 0) \qquad \forall n = 0, \cdots, N.$$
4. Définir l'enveloppe de Snell $(Z_k)_{k \geq 1}$ du processus $(Y_k)_{k \geq 1}$.
5. Montrer que pour tout $1 \leq k \leq N-1$, il existe une fonction mesurable ϕ_k tel que $Z_k = \phi_k(M_k - S_k)$ et
$$\phi_k(x) = \begin{cases} \max(p_{N-k},\ \frac{1}{2}(\phi_{k+1}(0) + \phi_{k+1}(1)) & \text{si } x = 0 \\ \frac{1}{2}(\phi_{k+1}(x+1) + \phi_{k+1}(x-1)) & \text{si } x \neq 0 \end{cases}$$
6. Pour tout entier $z \geq 0$ et tout $0 \leq n \leq N$, on définit
$$q(n,z) = \mathbb{P}(S_n = M_n, M_n \geq z).$$
Montrer que pour tout $0 \leq n \leq N$, $p_n = q(n,0)$ puis
$$\phi_k(x) = q(N-k, x) \quad \forall x \geq 0, \forall\ 0 \leq k \leq N$$
Indication : on pourra établir une relation de récurrence vérifiée par $q(\cdot,\cdot)$ *en appliquant la propriété de Markov à la marche au temps* 1.
7. En déduire un temps optimal.