

Correction détaillée de l'examen

Arrêt optimal et contrôle stochastique — 9 mai 2025

Objectif du document. Cette correction reprend l'énoncé fourni et donne une solution rédigée, en insistant sur les idées de programmation dynamique, l'enveloppe de Snell et l'interprétation économique des contrôles optimaux.

Exercice 1 — Arrêt optimal

On considère le payoff américain

$$Z_n = \varphi_n(X_n), \quad n = 0, \dots, N,$$

où (X_n) est une chaîne de Markov (inhomogène) et où l'on cherche

$$V_0 = \sup_{\tau \in \mathcal{T}_0} \mathbb{E}[Z_\tau].$$

Pour $n \in \{0, \dots, N\}$, on note $\mathcal{T}_n$ l'ensemble des temps d'arrêt à valeurs dans $\{n, \dots, N\}$.

1. Relation de récurrence rétrograde pour les fonctions valeurs

La fonction valeur à partir du temps n est

$$V_n(x) = \sup_{\tau \in \mathcal{T}_n} \mathbb{E}[Z_\tau \mid X_n = x].$$

Comme le modèle est markovien, les fonctions valeurs dépendent seulement de l'état courant x .

La programmation dynamique donne :

$$V_N(x) = \varphi_N(x),$$

et pour $n = N - 1, \dots, 0$,

$$V_n(x) = \max \left(\varphi_n(x), P_{n+1} V_{n+1}(x) \right),$$

où

$$P_{n+1} V_{n+1}(x) = \int_{\mathbb{R}_+^d} V_{n+1}(y) P_{n+1}(x, dy).$$

Autrement dit,

$$V_n(x) = \max \left(\varphi_n(x), \mathbb{E}[V_{n+1}(X_{n+1}) \mid X_n = x] \right).$$

Explication. À la date n , deux décisions seulement sont possibles :

- **exercer immédiatement**, ce qui rapporte $\varphi_n(x)$;
- **attendre au moins une période**, ce qui vaut en moyenne la valeur de continuation $P_{n+1} V_{n+1}(x)$.

On choisit donc le maximum des deux.

2. Que dire du processus $S_n = V_n(X_n)$?

Le processus

$$S_n = V_n(X_n), \quad n = 0, \dots, N,$$

est l'**enveloppe de Snell** du processus de gains (Z_n) .

En particulier :

1. (S_n) est adapté à la filtration $(\mathcal{F}_n)$;
2. il domine le payoff :

$$S_n \geq Z_n \quad \text{p.s. pour tout } n;$$

3. c'est une **surmartingale** :

$$\mathbb{E}[S_{n+1} \mid \mathcal{F}_n] \leq S_n;$$

4. c'est la **plus petite surmartingale dominante** de (Z_n) ;
5. le temps d'arrêt

$$\tau_n = \min\{k \geq n : S_k = Z_k\}$$

est optimal à partir de la date n ;

6. le processus arrêté $(S_{k \wedge \tau_n})_{k \geq n}$ est une martingale.

Pourquoi ? Par définition,

$$S_n = V_n(X_n) = \operatorname{ess\,sup}_{\tau \in \mathcal{T}_n} \mathbb{E}[Z_\tau \mid \mathcal{F}_n].$$

Cette écriture implique immédiatement $S_n \geq Z_n$ (on peut choisir $\tau = n$), et la propriété de surmartingale provient de la relation de Bellman.

3. Justification de la formule de continuation

On définit

$$C_N = 0, \quad C_n(x) = P_{n+1}V_{n+1}(x) = \mathbb{E}[V_{n+1}(X_{n+1}) \mid X_n = x], \quad n = 0, \dots, N-1.$$

On veut justifier que

$$C_n(x) = \mathbb{E}[Z_{\tau_{n+1}} \mid X_n = x].$$

L'idée est la suivante : à partir du temps $n+1$, le temps τ_{n+1} est optimal, donc

$$V_{n+1}(X_{n+1}) = \mathbb{E}[Z_{\tau_{n+1}} \mid \mathcal{F}_{n+1}].$$

En prenant l'espérance conditionnelle sachant $X_n = x$ et en utilisant la propriété de tour :

$$\begin{aligned} C_n(x) &= \mathbb{E}[V_{n+1}(X_{n+1}) \mid X_n = x] \\ &= \mathbb{E}[\mathbb{E}[Z_{\tau_{n+1}} \mid \mathcal{F}_{n+1}] \mid X_n = x] \\ &= \mathbb{E}[Z_{\tau_{n+1}} \mid X_n = x]. \end{aligned}$$

Lecture intuitive. La continuation à la date n est simplement la valeur espérée du payoff que l'on touchera si l'on suit ensuite la règle optimale à partir de $n+1$.

4. Relation de récurrence sur les temps d'arrêt τ_n

Par définition du critère optimal, on s'arrête à la date n si le gain immédiat est au moins aussi bon que la continuation. Or

$$V_n(X_n) = \max \left(Z_n, \mathbb{E}[Z_{\tau_{n+1}} \mid X_n] \right).$$

Donc

$$\tau_N = N,$$

et pour $n = N - 1, \dots, 0$,

$$\tau_n = \begin{cases} n, & \text{si } Z_n \geq \mathbb{E}[Z_{\tau_{n+1}} \mid X_n], \\ \tau_{n+1}, & \text{sinon.} \end{cases}$$

Si l'on pose

$$B_n = \mathbf{1}_{\{Z_n \geq \mathbb{E}[Z_{\tau_{n+1}} \mid X_n]\}},$$

on peut écrire de manière compacte

$$\tau_n = nB_n + \tau_{n+1}(1 - B_n).$$

Enfin, comme $X_0 = x_0$ est déterministe,

$$V_0(x_0) = \mathbb{E}[Z_{\tau_0}] = \max \left(\varphi_0(x_0), \mathbb{E}[Z_{\tau_1}] \right).$$

Interprétation. La région d'arrêt à la date n est

$$\mathcal{S}_n = \{x : \varphi_n(x) \geq C_n(x)\},$$

et la région de continuation est son complémentaire. Le temps optimal est le premier instant où l'état entre dans la région d'arrêt.

5. Algorithme de Longstaff–Schwartz

Le but est de construire numériquement une règle d'arrêt approchée à partir d'un échantillon simulé de trajectoires

$$(X_n^{(j)})_{n=0, \dots, N}, \quad j = 1, \dots, M.$$

On choisit au préalable une famille de fonctions de base $\psi_1, \dots, \psi_p$ sur l'espace d'état (polynômes, fonctions locales, etc.).

Étape 1 : simulation des trajectoires. Simuler M trajectoires indépendantes du processus (X_n) , puis calculer les payoffs

$$Z_n^{(j)} = \varphi_n(X_n^{(j)}).$$

Étape 2 : condition terminale. À maturité, il n'y a plus de choix :

$$\hat{\tau}_N^{(j)} = N, \quad j = 1, \dots, M.$$

Étape 3 : régression rétrograde. Pour $n = N - 1, \dots, 1$ (ou jusqu'à 0 si l'on souhaite tout traiter uniformément) :

- a) Pour chaque trajectoire j , calculer la valeur future réalisée sous la règle déjà construite :

$$Y_n^{(j)} = Z_{\hat{\tau}_{n+1}^{(j)}}^{(j)}.$$

- b) Effectuer une régression des variables $Y_n^{(j)}$ sur les fonctions de base évaluées en $X_n^{(j)}$:

$$Y_n^{(j)} \approx \beta_{n,1}\psi_1(X_n^{(j)}) + \dots + \beta_{n,p}\psi_p(X_n^{(j)}).$$

On obtient ainsi une approximation de la continuation :

$$\hat{C}_n(x) = \hat{\beta}_{n,1}\psi_1(x) + \dots + \hat{\beta}_{n,p}\psi_p(x).$$

- c) Définir le temps d'arrêt approché par la règle de décision

$$\hat{\tau}_n^{(j)} = \begin{cases} n, & \text{si } Z_n^{(j)} \geq \hat{C}_n(X_n^{(j)}), \\ \hat{\tau}_{n+1}^{(j)}, & \text{sinon.} \end{cases}$$

Étape 4 : estimateur de la valeur. Comme $X_0 = x_0$ est déterministe, on peut prendre par exemple

$$\hat{V}_0 = \max \left(\varphi_0(x_0), \frac{1}{M} \sum_{j=1}^M Z_{\hat{\tau}_1^{(j)}}^{(j)} \right),$$

ou, de façon équivalente si l'on a construit $\hat{\tau}_0^{(j)}$,

$$\hat{V}_0 = \frac{1}{M} \sum_{j=1}^M Z_{\hat{\tau}_0^{(j)}}^{(j)}.$$

Discussion sur les erreurs. Plusieurs sources d'erreur interviennent :

1. **Erreur de Monte Carlo** : due au nombre fini de trajectoires, typiquement d'ordre $M^{-1/2}$.
2. **Erreur de projection / approximation** : la vraie continuation C_n n'appartient pas en général à l'espace engendré par les bases $\psi_1, \dots, \psi_p$.
3. **Erreur statistique de régression** : même si l'espace de base est pertinent, les coefficients sont estimés avec un échantillon fini.
4. **Biais d'évaluation en-sample** : si l'on construit la politique et qu'on l'évalue sur les mêmes trajectoires, on peut introduire un biais de regard vers l'avant (*look-ahead bias*). En pratique, on préfère souvent évaluer la politique apprise sur un nouvel échantillon indépendant.

Remarque. La politique d'arrêt obtenue par Longstaff–Schwartz est en général sous-optimale. Son évaluation sur des trajectoires indépendantes fournit donc naturellement une **borne inférieure** du vrai prix américain.

Exercice 2 — Arbitrage épargne / consommation

L'énoncé définit la dynamique contrôlée

$$X_0 = x_0 > 0, \quad X_{n+1} = X_n + Y_{n+1}X_n(1 - U_n) = X_n(1 + Y_{n+1}(1 - U_n)),$$

où $U_n \in [0, 1]$ représente la part de richesse consommée à la date n , et (Y_n) est une suite i.i.d. positive, d'espérance finie

$$\mathbb{E}[Y_n] = \theta > 0.$$

On veut maximiser la consommation cumulée attendue. On comprend donc la fonctionnelle comme

$$\mathbb{E} \left[\sum_{k=0}^{N-1} X_k U_k \right],$$

(la somme écrite dans l'énoncé contient manifestement une petite coquille d'indice).

1. Définition du modèle markovien contrôlé

Le modèle markovien contrôlé est donné par :

- **Espace d'état** : $E = \mathbb{R}_+$;
- **Espace des actions** : $A = [0, 1]$;
- **Bruit** : Y_{n+1} i.i.d. sur $\mathbb{R}_+^*$;
- **Dynamique** :

$$X_{n+1} = F(X_n, U_n, Y_{n+1}), \quad F(x, u, y) = x(1 + y(1 - u));$$

- **Gain instantané** :

$$f_n(x, u) = xu;$$

- **Gain terminal** :

$$g(x) = 0.$$

Un contrôle admissible est une suite $(U_n)_{0 \leq n \leq N-1}$ telle que $U_n \in [0, 1]$ et U_n soit mesurable par rapport à l'information disponible à la date n .

2. Fonction valeur J_n et équations de programmation dynamique

On définit la valeur maximale à partir de la date n et de l'état x par

$$J_n(x) = \sup_{(U_k)_{k=n}^{N-1}} \mathbb{E} \left[\sum_{k=n}^{N-1} X_k U_k \mid X_n = x \right].$$

À l'horizon terminal,

$$J_N(x) = 0.$$

Pour $n = N - 1, \dots, 0$, l'équation de Bellman est

$$J_n(x) = \sup_{u \in [0, 1]} \left\{ xu + \mathbb{E} \left[J_{n+1}(x(1 + Y_{n+1}(1 - u))) \right] \right\}.$$

Explication. À la date n , on consomme d'abord xu , puis on laisse le reste fructifier aléatoirement, ce qui produit la richesse future X_{n+1} .

3. Forme linéaire de J_n et contrôles optimaux

On cherche une solution de la forme

$$J_n(x) = \rho_n x.$$

C'est très naturel ici car :

- la dynamique est **linéaire** en x ;
- le gain instantané xu est lui aussi **linéaire** en x .

En injectant cette forme dans Bellman :

$$\begin{aligned} J_n(x) &= \sup_{u \in [0,1]} \{xu + \rho_{n+1} \mathbb{E}[x(1 + Y_{n+1}(1 - u))]\} \\ &= x \sup_{u \in [0,1]} \{u + \rho_{n+1}(1 + \theta(1 - u))\}. \end{aligned}$$

Donc

$$\rho_N = 0,$$

et pour $n = N - 1, \dots, 0$,

$$\rho_n = \sup_{u \in [0,1]} \{u + \rho_{n+1}(1 + \theta(1 - u))\}.$$

Or l'expression à maximiser est affine en u :

$$u + \rho_{n+1}(1 + \theta(1 - u)) = \rho_{n+1}(1 + \theta) + u(1 - \theta\rho_{n+1}).$$

Le contrôle optimal est donc toujours **bang-bang** :

$$U_n^* = \begin{cases} 1, & \text{si } \rho_{n+1} < \frac{1}{\theta}, \\ 0, & \text{si } \rho_{n+1} > \frac{1}{\theta}, \\ \text{n'importe quel } u \in [0, 1], & \text{si } \rho_{n+1} = \frac{1}{\theta}. \end{cases}$$

En conséquence,

$$\rho_n = \max(1 + \rho_{n+1}, (1 + \theta)\rho_{n+1}).$$

Lecture économique.

- Si l'on consomme tout maintenant ($u = 1$), on gagne immédiatement x et la richesse reste ensuite constante : la valeur marginale devient $1 + \rho_{n+1}$.
- Si l'on ne consomme rien ($u = 0$), toute la richesse est investie : la valeur marginale devient $(1 + \theta)\rho_{n+1}$.

On compare donc *consommer maintenant* à *faire croître la richesse pour consommer plus tard*.

4. Existence d'une date de bascule unique K

Comme la suite (ρ_n) est croissante quand on remonte dans le temps (on ajoute des possibilités de consommation), le critère

$$\rho_{n+1} \gtrless \frac{1}{\theta}$$

ne peut changer de signe qu'une seule fois. Il existe donc une date de bascule $K \in \{0, \dots, N\}$ telle que la stratégie optimale soit :

$$U_n^* = 0 \quad \text{pour } n < K, \quad U_n^* = 1 \quad \text{pour } n \geq K,$$

avec la convention usuelle *consommer à l'égalité* lorsque $\rho_{n+1} = 1/\theta$.

Structure explicite de la valeur. À partir du moment où l'on commence à consommer tout, la richesse ne croît plus. Ainsi, pour $n \geq K$,

$$\rho_n = N - n.$$

En revanche, pour $n < K$, on n'investit rien en consommation et la valeur se multiplie par $(1 + \theta)$ à chaque pas :

$$\rho_n = (1 + \theta)^{K-n}(N - K).$$

En particulier,

$$J_0(x_0) = \rho_0 x_0 = (N - K)(1 + \theta)^K x_0.$$

Caractérisation de K . Au point de bascule, il faut que :

- à la date K , il soit optimal de **consommer** ;
- à la date $K - 1$, il soit optimal de **ne pas consommer**.

Comme, dans la zone de consommation, $\rho_K = N - K$ et $\rho_{K+1} = N - K - 1$, cela donne

$$\theta(N - K - 1) \leq 1 < \theta(N - K).$$

Autrement dit, $N - K$ est l'unique entier $m \in \{1, \dots, N\}$ tel que

$$m - 1 \leq \frac{1}{\theta} < m.$$

Avec la convention de consommation à l'égalité, on peut écrire

$$N - K = \min\left(N, \left\lfloor \frac{1}{\theta} \right\rfloor + 1\right), \quad K = N - \min\left(N, \left\lfloor \frac{1}{\theta} \right\rfloor + 1\right).$$

Interprétation économique. Le résultat est très intuitif :

- au début, s'il reste encore beaucoup de dates de consommation et que le rendement moyen θ est suffisamment attractif, il vaut mieux **épargner** ;
- à partir d'un certain instant, il n'y a plus assez de temps pour profiter de la capitalisation, donc il devient optimal de **consommer immédiatement et à 100%**.

La stratégie optimale est donc une stratégie de type *save first, consume later* avec une unique date de bascule.

Résumé ultra-court. Exercice 1 : on retrouve la programmation dynamique d'un problème d'arrêt optimal markovien, l'enveloppe de Snell et la règle *stop if immediate payoff $\geq$ continuation value*. Exercice 2 : la valeur est linéaire en la richesse, le contrôle optimal est bang-bang, et l'optimum consiste à ne rien consommer au début puis à tout consommer à partir d'une date seuil K .