Chapitre 2. Modèles à 1 période.

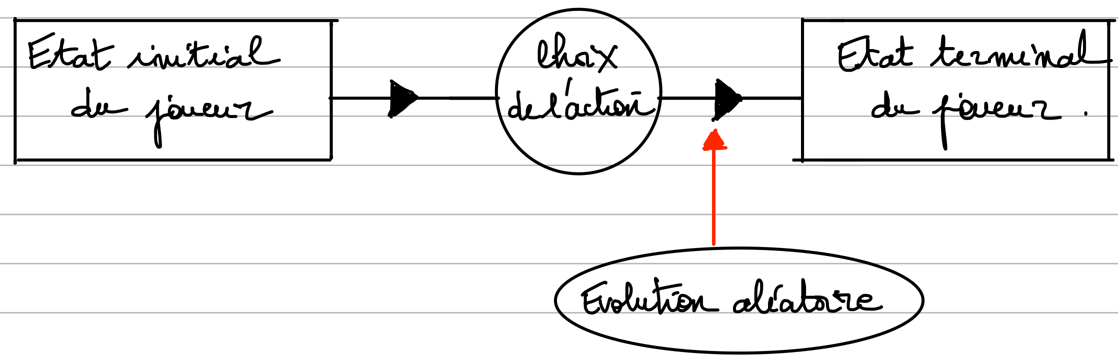
Dans ce chapitre, on considère un cas très particulier des processus de décision Markoviens : on suppose que le joueur n'a à choisir qu'une seule action avant que ne soit calculée sa récompense totale à l'issue du jeu.

Pour cette raison, on parle de jeux ou de modèles à 1 période.

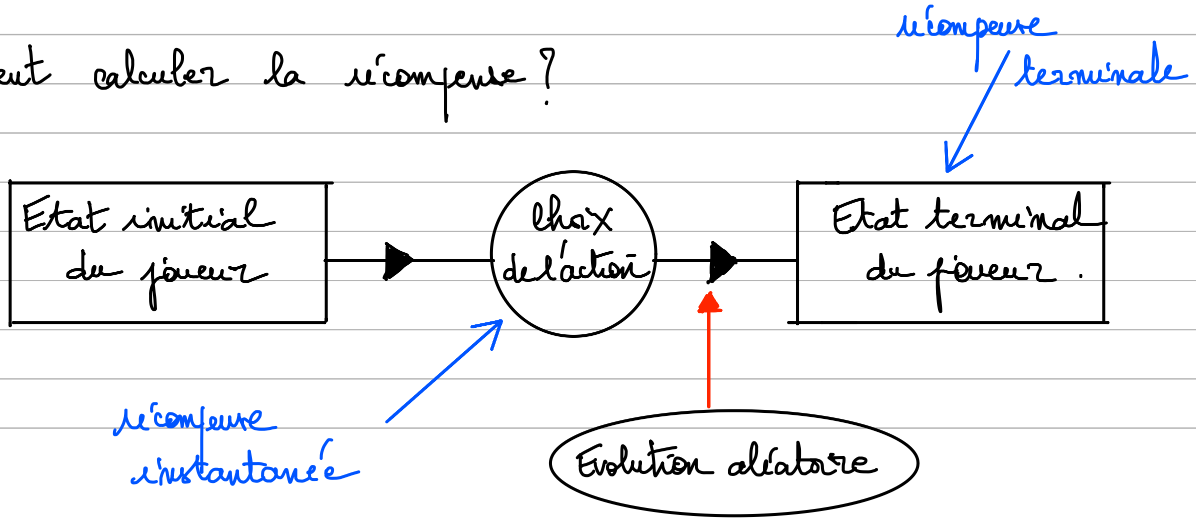
I Formalisation

1) Mise en place du cadre introduit dans le chapitre 1.

On résume le problème à l'aide d'une nouvelle figure



Comment calculer la récompense ?



On résume donc la récompense de la façon suivante :

Si • X_0 est l'état initial ($X_0 : \Omega \to S$)

• α_0 est l'action choisie par le joueur à l'instant 0 ($\alpha_0 : \Omega \to A$)

alors on appelle récompense instantanée la variable aléatoire

$$r(X_0, \alpha_0)$$

où r est une fonction mesurable

$$r : S \times A \longrightarrow \mathbb{R}$$
$$(s, a) \longmapsto r(s, a)$$

Une fois choisi A_0, l'état X_1 est obtenu en tirant un nouvel état selon $P(X_0, A_0, \cdot)$

on associe au nouvel état X_1 une nouvelle récompense

$$g(X_1)$$

où g est une fonction (nécessair^t mesurable)

$$g : S \longrightarrow \mathbb{R}$$
$$s \longmapsto g(s).$$

2) Caractère aléatoire de la récompense

Attention : La condition initiale peut être ALEATOIRE

— (adaptée à une tribu $\mathcal{F}_0$).

De même α_0 peut être aléatoire

Donc $r(X_0, A_0)$ est vue comme une variable aléatoire

Observation : Il est clair que X_1 peut être aléatoire —

—

Quelle est la récompense totale :

$$\underbrace{r(X_0, A_0) + g(X_1)}_{\text{il s'agit d'une variable aléatoire}}.$$

3) problème d'optimisation

Objectif ; On veut optimiser la récompense. Mais quel sens cela a-t-il ?

Première tentative : Pour chaque réalisation du 'hasard', on veut trouver une action α_0 maximisant la récompense. Autrement dit, pour $\omega \in \Omega$ (réalisation du hasard), on veut résoudre

$$\max_{\beta \in A} \left\{ r(X_0(\omega), \beta) + g(X_1(\omega)) \right\}$$

et poser ensuite

$$\alpha_0(\omega) \in \operatorname*{Argmax}_{\beta \in A} \left\{ r(X_0(\omega), \beta) + g(X_1(\omega)) \right\}$$

Ecrit, comme cela le problème n'est pas bien écrit car $X_1(\omega)$ dépend implicitement du choix de l'action. Pour rendre ceci plus évident, on rappelle la formulation canonique

$$X_1(\omega) = F_0^1 \left(X_0(\omega), \alpha_0(\omega), U_1(\omega) \right)$$

où $U_1 \sim \mathrm{Unif}\,[0,1]$ est indépendant de (X_0, A_0) (ce qui est vrai si U_1 est indépendant de $\mathcal{F}_0$) et

$$F_0 : S \times A \times [0,1] \longrightarrow S \quad \text{(mesurable)}$$

On devrait donc réécrire la maximisation sous la forme

$$\alpha_0(\omega) \in \operatorname*{Argmax}_{\beta \in A} \left\{ r(X_0(\omega), \beta) + F_0(X_0(\omega), \beta, U_1(\omega)) \right\}$$

Observations :

- Si l'ensemble A est fini, l'existence d'un maximiseur ne pose aucune difficulté.
- En revanche la mesurabilité de l'application

$$\omega \longmapsto \alpha_0(\omega)$$

n'est pas évidente. En effet, la construction de α_0 consiste à considérer, pour $x_0 \in S$ et $u \in [0,1]$,

$$a_0(x_0, u) \in \operatorname*{Argmax}_{\beta \in A} \left\{ r(x_0, \beta) + F(x_0, \beta, u) \right\}$$

et ensuite à poser

$$\alpha_0(\omega) = a_0(X_0(\omega), U_1(\omega))$$

Le problème en pratique est qu'il peut être difficile de montrer la mesurabilité de a_0 (en fait, cela ne pose pas de difficulté si le maximiseur est unique, mais la sélection est délicate en cas de multiplicité du maximiseur). La mesurabilité de a_0 est comprise au sens :

$$a_0 : S \times [0,1] \longrightarrow A$$

- En réalité, même si a_0 est mesurable, la solution donnée ici est mauvaise. Regardons en effet la mesurabilité de α_0

$$\alpha_0 = a_0(X_0, U_1)$$

si a_0 est mesurabilité alors α_0 N'EST PAS $\mathcal{F}_0$-mesurable car U_1 n'est pas $\mathcal{F}_0$-mesurable (sauf cas très particuliers sur la structure de a_0).

Cette solution est donc <u>mauvaise</u> !

Conséquence : Nous sommes contraints à maximiser — en moyenne (par opposition à une maximisation aléa par aléa).

On cherche donc à maximiser :

$$\mathbb{E}\{ r(X_0, \alpha_0) + g(X_1) \}$$

en choisissant la meilleure variable aléatoire $\mathcal{F}_0$ - mesurable :

$$\alpha_0 : \Omega \longrightarrow A$$

4) Résolution du problème d'optimisation.

On veut maintenant résoudre le problème identifié dans le § 3).

Cela suppose de préciser la valeur de l'$\mathbb{E}$, pour un choix donné de α_0.

$$\mathbb{E}\{ r(X_0, \alpha_0) + g(X_1) \}$$

$$= \mathbb{E}\{ \mathbb{E}[\underbrace{r(X_0, \alpha_0)}_{\mathcal{F}_0\text{ - mesurable}} + g(X_1) \mid \mathcal{F}_0] \}.$$

$$= \mathbb{E}\{ r(X_0, \alpha_0) + \underbrace{\mathbb{E}[g(X_1) \mid \mathcal{F}_0]} \}.$$

$$\sum_{j \in S} g(j)\ \underbrace{\mathbb{P}(\{X_1 = j\} \mid \mathcal{F}_0)}_{P(X_0, \alpha_0, j)}$$

$$= \mathbb{E}\{ r(X_0, \alpha_0) + \sum_{j \in S} g(j)\, P(X_0, \alpha_0, j) \}.$$

Cas 1 : Supposons que $\mathbb{P}\{X_0 = i\} = 1$
pour un certain $i \in S$.

Alors l'espérance s'écrit
$$\mathbb{E}\left\{ r(i, \alpha_0) + \sum_{j\in S} g(j)\, \mathbb{P}(i, \alpha_0, j) \right\}.$$

On en déduit

Lemme : Si $\mathbb{P}\{X_0 = i\} = 1$ pour un certain $i \in S$
et si l'application
$$a \in A \longmapsto r(i, a) + \sum_{j\in S} g(j)\, \mathbb{P}(i, a, j)$$
admet un maximiseur $a(i) \in A$
alors la v.a
$$\alpha_0(\omega) = a(i), \text{ pour } \omega \in \Omega,$$
est solution du problème.

Preuve : Il suffit de montrer que, pour toute autre variable aléatoire :
$$\beta_0 : \Omega \to A \quad \mathcal{F}_0\text{-mesurable},$$
on a
$$\mathbb{E}\left\{ r(i, \beta_0) + \sum_{j\in S} g(j)\, \mathbb{P}(i, \beta_0, j) \right\}$$
$$\leq \mathbb{E}\left\{ r(i, \alpha_0) + \sum_{j\in S} g(j)\, \mathbb{P}(i, \alpha_0, j) \right\}.$$

Pour tout $\omega \in \Omega$
$$r(i, \beta_0(\omega)) + \sum_{j\in S} g(j)\, \mathbb{P}(i, \beta_0, j)$$

$$\leq \max_{a \in A} \left\{ r(i,a) + \sum_{j \in S} g(j)\ P(i,a,j) \right\}$$

$$= r(i, a(i)) + \sum_{j \in S} g(j)\ P(i, a(i), j)$$

$$= r(i, \alpha_0(\omega)) + \sum_{j \in S} g(j)\ P(i, \alpha_0(\omega), j)$$

Le résultat suit en prenant l'espérance. $\square$.

Observation : Si A est fini, l'existence d'un maximiseur est évidente.

Cas 2 (général) :

Toute l'idée des processus de décision Markoviens à 1 période réside dans le fait que la solution précédente reste vraie si X_0 est aléatoire

Proposition : Supposons que, pour tout $i \in S$, l'application

$$a \in A \longmapsto r(i,a) + \sum_{j \in S} g(j)\ P(i,a,j)$$

admet un maximiseur $a(i) \in A$

alors la v.a

$$\boxed{\alpha_0(\omega) = a(X_0(\omega)),\ \omega \in \Omega}$$

est solution du problème de maximisation.

Preuve : On note que $a : S \ni i \mapsto a(i) \in A$ est nécessairement mesurable car S est dénombrable.

On reprend maintenant la preuve du lemme, que l'on applique sur chaque événement de la forme

$$\{X_0 = i\},\quad i \in S.$$

La preuve donne pour $\omega \in \{X_0 = a(i)\}$

$$r(X_0(\omega), \beta_0(\omega)) + \sum_{j \in S} g(j)\, P(X_0(\omega), \beta_0(\omega), j)$$

$$\leq r(X_0(\omega), a(i)) + \sum_{j \in S} g(j)\, P(X_0(\omega), a(i), j)$$

$$= r(X_0(\omega), a(X_0(\omega))) + \sum_{j \in S} g(j)\, P(X_0(\omega), a(X_0(\omega)), j).$$

Cette inégalité est vraie pour chaque $i \in S$, donc sur Ω. On prend l'$\mathbb{E}$ et on conclut. $\square$.

Remarque : La solution trouvée est automatiquement sous forme Markovienne : elle est au moins aussi bonne que toute autre solution dans laquelle l'action jouée incorpore un aléa supplémentaire en plus de celui porté par X_0.

Exercice (TD) : Reprendre le modèle à $|S| = 2$, $|A| = 2$ du chapitre 1 et expliciter la/les stratégie(s) optimale(s) pour une fonction

$$g : \{0,1\} \to \mathbb{R}.$$

II. Exemple : allocation optimale sur un marché financier à une période

1) Modélisation du marché

Deux actifs financiers :

- Un cours initial (à $t=0$)
  final (à $t=1$)

- Un actif sans risque

$$S_0^0 = 1 \quad \text{(par simplicité)}$$

$$S_1^0 = 1+r \quad \text{(où } r \text{ est un rendement sûr)}$$

- Un actif risqué

S_0, valeur positive déterministe

$S_1 = S_0 \xi$ où ξ est un facteur aléatoire.

Définition de la filtration

$$\mathcal{F}_0 = \{\phi, \Omega\}.$$

$$\mathcal{F}_1 = \sigma(\xi),$$

2) Richesse

La variable d'état représente la richesse d'un investisseur.

W_0 : capital initial à investir par un investisseur.

Quelle est la variable de contrôle ?

<u>Réponse</u> : La stratégie d'investissement

<u>Définition</u> : On appelle stratégie d'investissement un réel $\phi \in \mathbb{R}$ représentant le nombre (réel) de parts détenues par l'investisseur à l'instant 0.

<u>Remarque</u> : Si $\phi < 0$, l'investisseur est endetté dans l'actif risqué et doit "rembourser" à $t=1$ le nombre de parts empruntées à $t=0$

Variable d'état à $t=0$:

$$W_0 = \phi S_0 + \underbrace{\frac{W_0 - \phi S_0}{S_0^0}}_{\text{nombre de parts d'actif sans risque détenues à } t=0} S_0^0$$

$S_0^0 = 1.$

Variable d'état à $t=1$:

$$W_1 = \phi S_1 + \frac{W_0 - \phi S_0}{S_0^0} S_1^0$$

$$= \xi \phi S_0 + \frac{W_0 - \phi S_0}{S_0^0} (1+r) S_0^0$$

$S_0^0 = 1.$

Exercice (TD) : Réplicabilité dans le modèle de Cox.

On appelle modèle de Cox le modèle dans lequel

$$\xi : \Omega \to \{d, u\} \quad \text{où} \quad 0 < d < u,$$

avec $\mathbb{P}\{\xi = u\} \in]0, 1[$.

On dit que le modèle est réplicable si, pour toute fonction

$$f : \mathbb{R} \to \mathbb{R}$$

il existe une richesse initiale $W_0 \in \mathbb{R}$ et une stratégie $\phi \in \mathbb{R}$ telle que

$$\mathbb{P}\{\underbrace{W_1^\phi}_{\text{richesse dirigée par } \phi} = f(\xi)\} = 1.$$

Montrer que, sous l'hypothèse

$$d < 1+r < u$$

le modèle est réplicable !

3) Optimisation.

Définition : On appelle fonction d'utilité toute fonction

$$U : I \longrightarrow \mathbb{R}$$

définie sur un intervalle I telle que

1) U continue sur I

2) U strict^t croissante sur I

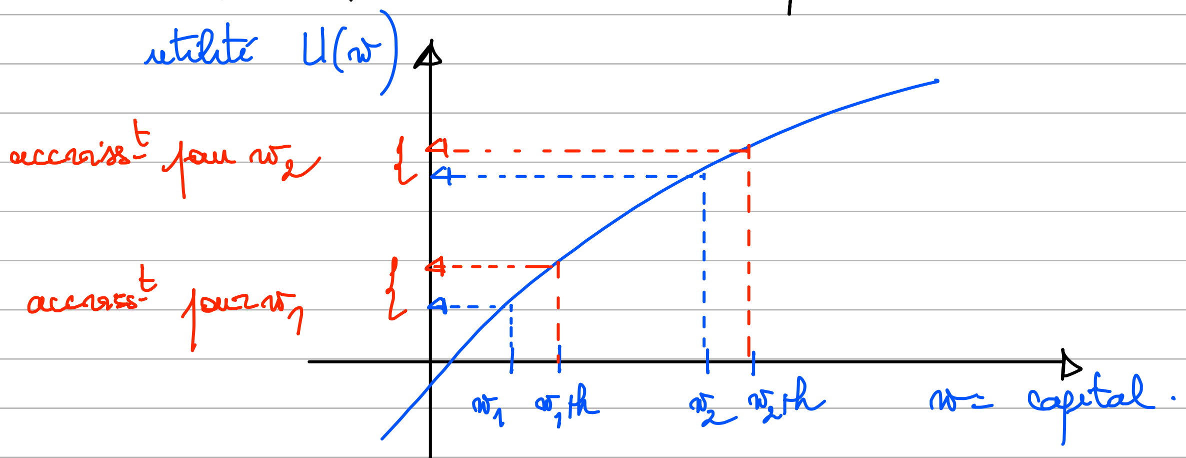
3) U strict^t concave sur I

$$\forall\, x, x' \in I \qquad U(\lambda x + (1-\lambda)x')$$
$$\lambda \in]0,1[\qquad > \lambda\, U(x) + (1-\lambda)\, U(x').$$

En général, on calcule $U(w)$ pour w richesse :

1) la satisfaction est d'autant plus grande que la richesse est élevée

2) pour la concavité, faisons un dessin



L'accroiss^t de l'utilité est supérieur pour w_1 : pour obtenir un accroiss^t donné, il est nécessaire d'investir davantage si la richesse initiale est plus élevée.

Exemples (TD) :

i) utilités exponentielles

$$U(x) = -\exp(-\gamma x), \quad x \in \mathbb{R},$$

où $\gamma > 0$

ii) utilités puissances

$$U(x) = x^{\gamma}, \quad x \geq 0,$$

où $\gamma \in]0,1[$

$$U(x) = -x^{\gamma} \quad x > 0,$$

où $\gamma < 0$.

iii) utilité logarithmique

$$U(x) = \ln(x) \quad x > 0.$$