Les problèmes d'arrêt optimal et leur résolution

Mémoire de Master en Mathématiques

Option: Analyse Mathématiques et Application

Présenté par : HASSANINE Chérifa

Encadré par : Mr. A. KALI

Université de Djilali Bounaâma-Khemis miliana

6 Juillet 2019

Plan

Plan

Plan

Plan

Plan

Plan

- La théorie **des problèmes d'arrêt optimal** consiste à déterminer les instants d'arrêt optimal opportun pour effectuer une action particulière, afin de maximiser la récompense attendue ou de minimiser le coût attendu.
- **L'intérêt d'un problème d'arrêt optimal** vient du fait qu'une vaste classe de problèmes peut être modélisés en utilisant cette formalisme : les applications statistiques qui exigent l'arrêt des observations à un certain moment.

Définition

Un processus stochastique $\mathbb{X} = (X_t)_{t \in T}$ est une famille de variables aléatoires X_t indexée par un ensemble T.

Dans notre cas, les processus stochastiques sont considérés en temps et espace d'états Ω discrets.

Définition

La probabilité de transition de l'état X_t à l'état X_{t+1} est :

$$P(X_{t+1}/X_t, X_{t-1}, \ldots, X_0)$$

Propriété de Markov

Un processus stochastique vérifiant **la propriété de Markov** si :

$$P(X_{t+1}/X_t, X_{t-1}, \ldots, X_0) = P(X_{t+1}/X_t), \quad \forall t \in T$$

Cette dernière est appelé un processus de Markov ou processus markovien.

Définition

On dit que $(X_n)_{n\in\mathbb{N}}$ est **une chaîne de Markov** si :

$$P\left(X_{n+1}=j|X_n=i,X_{n-1}=i_{n-1},\ldots,X_1=i_1,X_0=i_0\right)=$$

$$P\left(X_{n+1}=j|X_n=i\right)$$

La chaîne de Markov est finie si l'espace d'état Ω est fini.

Définition

La matrice de transition d'une chaîne de Markov est la matrice carrée $P = (m_{ij})$ dont le coefficient m_{ij} est la probabilité de transition P_{ij} d'état i a l'état j.

$$P = \begin{bmatrix} p_{0,0} & p_{0,1} & \cdots \\ p_{1,0} & p_{1,1} & \cdots \\ \vdots & \vdots & \vdots \end{bmatrix}$$

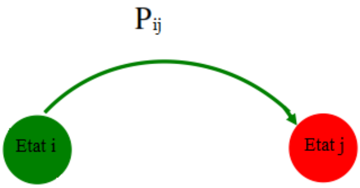


FIGURE – Graphe représentatif de la chaine de Markov

Définition

Processus Décisionnels de Markov est défini par un quatuple (Ω, A, P, R) ou :

- Ω est l'espace d'états du processus qui peut être fini, ou dénombrable.
- A est l'ensemble des décisions ou d'actions a entreprendre.
- P est une matrice de transition entre les états du processus.
- R est une fonction de coût.

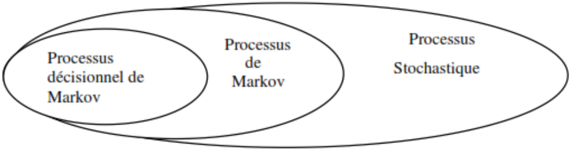


FIGURE – Structure des processus stochastiques

- Résoudre un problèmes de décision consiste à trouver la meilleur politique dit **optimale**.

  Pour certain critère, en général trois critères sont à considérer.

Critères d'optimalité :

Pour toute politique π, on définit :

- Critères de l'espérance du coût prévisionnel

$$\mathbb{E}_\pi \left\{ \sum_{t=0}^{\infty} \alpha^t R(x(t)) \mid x(0) = i \right\} ; i \geq 0$$

- Critères de l'espérance du coût positif non prévisionnel

$$\mathbb{E}_\pi \left\{ \sum_{t=0}^{\infty} R(x(t)) \mid x(0) = i \right\} ; i \geq 0$$

- Critères de l'espérance du coût moyen

$$\lim_{n \to +\infty} \frac{\mathbb{E}_\pi \left\{ \sum_{t=0}^{n} \alpha^t R(x(t)) \mid x(0) = i \right\}}{n+1} ; i \geq 0$$

- Résoudre un problème de décision consiste à trouver la meilleure politique, c'est-à-dire celle **maximise le critère de l'espérance du coût prévisionnel** pour chaque état initial i.

  Posons :

$$V_\alpha(i) = \sup_\pi V_\pi(i), \quad i \geq 0$$

- **La fonction de gain optimal** V_α est donc :

$$V_\alpha(i) = \sup_\pi \mathbb{E}_\pi \left[\sum_{t=1}^{\infty} \alpha^t R(X(t), a(t)) | X(0) = i \right], i \geq 0$$

Théorème

Pour tout état i d'un processus décisionnel de Markov, on a :

$$V_\alpha(i) = \max_a \left\{ R(i,a) + \alpha \sum_{j=0}^{\infty} P_{ij}(a) V_\alpha(j) \right\}, i \geq 0$$

- **Une politique stationnaire** est une fonction de l'espace d'états dans l'espace des décisions.
- **Une politique déterministe** associe à chaque état une et une seule action. Elle s'oppose aux politiques stochastiques.

- Pour tout PDM possédait au moins une politique optimale déterministe et stationnaire. De ce fait, la recherche de la politique peut être réalisée dans l'espace des politiques déterministes.
- Une politique f est dite stationnaire si l'action $f(i)$ choisie à l'état i et à l'instant t dépend seulement de l'état du processus en cet instant.

Définition

Un problème d'arrêt optimal pour une chaîne de Markov est défini par le cinqtuple $(\Omega, P, R, M, \alpha)$ tel que :

- Ω l'ensemble d'états.
- P une matrice de transition entre les états de la chaîne. $P_{ij}(a)$ représente la probabilité de passer de l'état i à l'état j.
- R une fonction de gain.
- M une fonction gain terminal.
- $\alpha \in]0, 1[$ un nombre réel, appelé taux d'actualisation.

- **l'objectif** celui de **maximiser** l'espérance des coûts linéaires et prévisionnels $V_\pi(i)$ sur tout les instant d'arrêt t :

$$V_\pi(i) = \mathbb{E}_\pi \left\{ \sum_{t=0}^{t-1} -\alpha^t R(X(t)) + \alpha^t M(i) \right\}$$

- La fonction profit maximale $V_\alpha(i)$ du le processus X qu'on dénoter $V_\alpha(i, M(i))$ est donnée par :

$$V_\alpha(i, M(i)) = \max_a \left\{ -R(i) + \alpha \sum_{j=0}^{N} P_{ij}(a) V_\alpha(j); M(i) \right\}$$

Méthodes de résolution des problèmes d'arrêt optimal :

- La programmation dynamique.
- la méthode d'itération sur les valeurs, pour résoudre La fonction profit maximale $V_\alpha(i)$. On considère une suite de fonction $V_n(i, M(i))$ définie par :

$$\begin{cases} V_{n+1}(i, M(i)) = \max \left[M(i); R(i) + \alpha \sum_{j=0}^{N} P_{ij} V_n(j, M(i)) \right]; \\ V_0(i, M(i)) = M(i) \end{cases}$$

Dans ce cas $V_0(i, M(i)) \leq V_1(i, M(i)) \leq \ldots$, la suite $V_n(i, M(i))$ converge vers $V(i, M(i))$.

- L'équation de Bellman quand l'espace d'état est fini.
- Davis et Karatzas(1994). Ils fournissent une interprétation intéressante de la décomposition de Doob Mayer de l'enveloppe de Snell.
- L'algorithme d'élimination de SONIN.

Les avantages et les inconvénients :

- La première méthode est la meilleur dans tous les cas où elle peut être appliquée mais l'inconvénient qu'il existe des limitations évidentes à son application.
- L'avantage de la deuxième méthode est qu'elle est universelle, mais la convergence de V_n vers V peut être très lente même dans un espace d'état avec quelques points s'il y a un cycle à suivre avec une probabilité proche de un.

- La troisième méthode est difficile à utiliser pour des butes de calcul.
- La cinquième méthode est elle résout facilement des exemples en utilisant quelques calculs simples.

- **L'algorithme d'élimination de SONIN** est basé sur la considération suivante : Dans un problème d'arrêt optimal, il est difficile de trouver les états où il est optimal, mais il est relativement facile de trouver un état ou des états où il n'est pas optimal d'arrêter.
- **l'objectif** de cette algorithme est déterminer les v(i) où :

$$V(i) = \max \left\{ -R(i) + \alpha \sum_{j=0}^{N} P_{ij} V(j); M(i) \right\}$$

Algorithm 1 : ALGORITHME DE SONIN

- **Input :**

  L'espace d'états Ω_1.

  La valeur du taux d'actualisation α dans l'intervalle $]0, 1[$.

  Les valeurs des probabilités de transition $P_0(i, j), \quad \forall i, j \in \Omega_1$.

  Les valeurs des gains $R_1(i), \quad \forall i \in \Omega_1$.

  Les valeurs des gains terminales $M(i), \quad \forall i \in \Omega_1$.

- **Output :**

  Les valeurs des $V(i), \quad \forall i \in \Omega_1$.

- **Étape 1**

  Ajouter à l'espace d'états Ω_1 un état absorbant x^*.

  Recalculer les probabilités de transition

  $P_1(i,j), \quad \forall i,j \in \Omega_1 \cup \{x^*\}$, avec :

  1. $P_1(i,j) = \alpha P_0(i,j), \quad \forall i,j \in \Omega_1$
  2. $P_1(i,x^*) = 1-\alpha, \quad \forall i \in \Omega_1$
  3. $P_1(x*,j) = 0, \quad \forall j \in \Omega_1$
  4. $P_1(x^*,x^*) = 1$

- Calculer $d(i) = M(i) - \left\{ R_1(i) + \sum_{j \in \Omega_1} P_1(i,j) M(j) \right\}, \quad \forall i \in \Omega_1$
- déterminer l'ensemble $D_1 = \{ i \in \Omega_1 / d(i) < 0 \}$

Si $D_1 = \phi$ faire

$V(i) = M(i), \quad \forall i \in \Omega_1$

Sinon

Si $D_1 = \Omega_1$ faire

"il n'est pas optimal de s'arrêter à n'importe quel état du processus."

Sinon

- **Étape 2 :**

$f := 1$

Tant que $D_f \neq \phi$ faire

$$\Omega_{f+1} = \Omega_f - D_f$$

Si $\mathbf{D_f} = \{\mathbf{k}\}$ (D_f contient un seul élément noté k) faire

$$P_{f+1}(i,j) = P_f(i,j) + \frac{P_f(i,k)P_f(k,j)}{1 - P_f(k,k)}, \quad \forall i,j \in \Omega_{f+1} \cup \{x*\}$$

$$R_{f+1}(i) = R_f(i) + \frac{P_f(i,k)R_f(k)}{1 - P_f(k,k)}, \quad \forall i \in \Omega_{f+1}$$ voir ici

Sinon (D_f contient plus d'un élément)

$$P'_f := \{P_f(i,j)\}, \quad i\,, j \in \Omega_{f+1}$$

$$H_f := \{P_f(i,j)\}, \quad i \in \Omega_{f+1} \text{ et } j \in D_f$$

$$Q_f := \{P_f(i,j)\}, \quad i,j \in D_f$$

$$T_f := \{P_f(i,j)\}, \quad i \in D_f \text{ et } j \in \Omega_{f+1}$$

$$N_f = (1 - Q_f)^{-1}$$

$$P_{f+1} = P_f + H_f N_f T_f$$

$$R_{f,D_f} = \{R_f(i)\}\,, \quad \forall i \in D_f$$

$$R_{f,\Omega_{f+1}} = \{R_f(i)\}\,, \quad \forall i \in \Omega_{f+1}$$

$$R_{f+1} = R_{f,\Omega_{f+1}} + H_f N_f R_{f,D_f}$$

voir ici

- Calculer
$$d(i) = M(i) - \left[R_{f+1}(i) + \sum_{j \in \Omega_{f+1}} P_{f+1}(i,j) M(j) \right], \forall i \in \Omega_{f+1}$$
- Déterminer l'ensemble $D_{f+1} = \{ i \in \Omega_{f+1} / d(\mathrm{i}) < 0 \}$

$\Omega_f = \Omega_{f+1}$;

$D_f = D_{f+1}$;

Fin tant que

Si $D_{f+1} = \phi$

$V(i) = M(i), \quad \forall i \in \Omega_f$;

$$V(i) = M(i) - \left[R_{f+1}(i) + \sum_{j \in \Omega_f} P_{f+1}(i,j) V(j) \right], \quad \forall i \in D_f$$

L'exécution du programme qu'on a implémenté voir ici

Théorème de Kolmogorov, Doblin

1. $P_2 = P_1' + H_1 U_1 = P_1' + H_1 N_1 T_1$ avec $P_1 = \begin{bmatrix} Q_1 & T_1 \\ H_1 & P_1 \end{bmatrix}$

   Où $P_1' = \{P_1(i,j)\}\,;\, i,j \in \Omega_2$

   $H_1 = \{P_1(i,j)\}\,;\, i \in \Omega_2$ et $j \in D$

   $Q_1 = \{P_1(i,j)\}\,;\, i,j \in D$

   $T_1 = \{P_1(i,j)\}\,;\, i \in D \text{et} j \in \Omega_2$

2. $R_2 = R_{1,\Omega 2} + H_1 N_1 R_{1,D}$
   Où $R_{1,\Omega 2} = \{R_1(i)\}\,;\, i \in \Omega_2$ et $R_{1,D} = \{R_1(i)\}\,;\, i \in D$.

   Quand $D = \{k\}$ on a :

3. $P_2(i,j) = P_1(i,j) + P_1(i,k)\,(1 - P_1(k,k))^{-1}\, P_1(k,j);\, i,j \in \Omega_2$

4. $R_2(i) = R_1(i) + P_1(i,k)\,(1 - P_1(k,k))^{-1}\, R_1(k);\, i \in \Omega_2$ retour

Exemple :

On considère un problème d'arrêt optimal pour une chaîne de Markov ou l'espace états est $\Omega = \{1, 2, 3\}$

- Le taux d'actualisation est $\alpha = 0.9$
- La matrice P_0 des probabilités de transition est :

$$P_0 = \begin{bmatrix} 0.33 & 0.33 & 0.34 \\ 0.5 & 0.25 & 0.25 \\ 0.33 & 0.5 & 0.17 \end{bmatrix}$$

- Le vecteur R_1 des gains est :

$$R_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

- Le vecteur M des gains terminaux est :

$$M = \begin{bmatrix} -1.5 \\ 2 \\ 4 \end{bmatrix}$$

Étape 1 :

- ajoutons à l'espace d'états Ω_1 un état absorbant x^*.

$$\Omega_1 \cup \{x^*\} = \{1, 2, 3, x^*\}$$

- la matrice des probabilités de transition P_1 pour le nouveau espace d'état $\Omega_1 \cup \{x^*\}$ est :

$$P_1 = \begin{bmatrix} 0.297 & 0.297 & 0.306 & 0.1 \\ 0.45 & 0.225 & 0.225 & 0.1 \\ 0.297 & 0.45 & 0.153 & 0.1 \\ 0. & 0. & 0. & 1 \end{bmatrix}$$

- Déterminons les valeurs de $d(i)$ telle que :

$$d(i) =: M(i) - \left[R_1(i) + \sum_{j \in \Omega_1} P_1(i,j) M(j) \right]$$

$$d_i = \begin{bmatrix} -3.8725 \\ 2.325 \\ 1.9335 \end{bmatrix}$$

- L'ensemble D_1 est $D_1 = \{i \in \Omega_1 / d(i) < 0\}$.

  D'où :

$$D_1 = \{1\}$$

- L'ensemble $\Omega_2 = \Omega_1 - D_1$.

  Ce qui donne :

$$\Omega_2 = \{2, 3\}$$

Étape 2 :

- la matrice des probabilités de transition P_2 pour $\{2, 3, x^*\}$ est :

$$P_2 = \begin{bmatrix} 0.415 & 0.42 & 0.164 \\ 0.575 & 0.282 & 0.142 \\ 0. & 0. & 1 \end{bmatrix}$$

- Le vecteur des gains R_2 pour $\Omega_2 = \{2, 3\}$ est :

$$R_2 = \begin{bmatrix} -0.359 \\ 1.422 \end{bmatrix}$$

- Pour tout état i de Ω_2, les $d(i)$ calculés par la formule

  $d(i) = M(\mathrm{i}) - \left[R_2(i) + \sum_{j \in \Omega_2} P_2(i,j) M(j) \right]$ sont donnés par :

  $d(2) = -0.153$

  $d(3) = 0.297$

- L'ensemble D_2 est $D_2 = \{i \in \Omega_1 / d(i) < 0\}$.

  D'où :

$$D_2 = \{1\}$$

- L'ensemble $\Omega_3 = \Omega_2 - D_1$.

  Ce qui donne :

$$\Omega_3 = \{3\}$$

- le principe de la deuxième étape est répété une fois, pour l'espace d'états $\Omega_3 = \{3\}$.
- Pour tout état i de Ω_3, les $d(i)$ calculés par la formule $d(i) = M(\mathrm{i}) - \left[R_3(i) + \sum_{j \in \Omega_3} P_3(i,j) M(j) \right]$ sont donnés par :
  $d(3) = 1,448$
  Qui est un nombre positif. Par conséquent, l'ensemble d'arrêt est obtenu c'est $\{3, x^*\}$.

- Pour l'état i :

$$V(3) = M(3) = 4$$

- Pour tout état i de D_2, $V(2)$ est calculé par la formule :

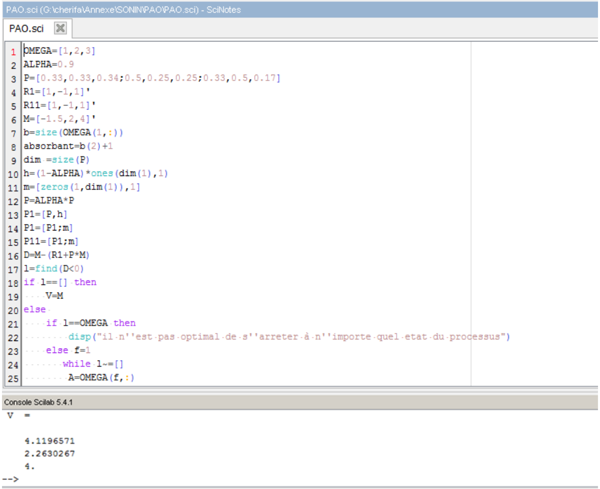
$$V(i) = \frac{1}{1 - p_3(i,i)} \left[R_3(i) + \sum_{j \in \Omega_3} P_3(i,j) V(j) \right]$$

D'où, $V(2) = 2.263$

- Pour tout état i de D_1, $V(1)$ est calculé par la formule :

$$V(i) = \frac{1}{1 - p_2(i,i)} \left[R_2(i) + \sum_{j \in \Omega_2} P_2(i,j) \mathrm{V}(j) \right]$$

D'où, $V(1) = 4.119$


PAO.sci (G:\cherifa\Annexe\SONIN\PAO\PAO.sci) - SciNotes

PAO.sci

```
OMEGA=[1,2,3]
ALPHA=0.9
P=[0.33,0.33,0.34;0.5,0.25,0.25;0.33,0.5,0.17]
R1=[1,-1,1]'
R11=[1,-1,1]'
M=[-1.5,2,4]'
b=size(OMEGA(1,:))
absorbant=b(2)+1
dim =size(P)
h=(1-ALPHA)*ones(dim(1),1)
m=[zeros(1,dim(1)),1]
P=ALPHA*P
P1=[P,h]
P1=[P1;m]
P11=[P1;m]
D=M-(R1+P*M)
l=find(D<0)
if l==[] then
    V=M
else
    if l==OMEGA then
        disp("il n''est pas optimal de s''arreter à n''importe quel etat du processus")
    else f=1
       while l~=[]
        A=OMEGA(f,:)
```

Console Scilab 5.4.1

```
 V  =

    4.1196571
    2.2630267
    4.
-->
```

FIGURE – Fenêtres de l'exécution du programme. retour

Dans cette présentation, nous avons présente **les problème d'arrêt optimal** pour une chaine de Markov a temps et espace d'état discret.

Plusieurs algorithmes sont proposes pour résoudre le problème optimale-ment. le plus reconnu est celui de Sonin.

Notre attention est portée sur la présentation et l'implémentation de cet algorithme.

Merci de votre attention !