Examen du module MODE

C. Gonzales / P. Perny

Durée : 3 heures

Seuls documents autorisés :
Les transparents de cours, les notes de cours ainsi que les articles étudiés dans la 2ème partie du cours sont les seuls documents autorisés. Les calculatrices seront également autorisées.

Exercice 1 (7 points – Processus décisionnels Markoviens)

Un décideur observe un système à événements discrets qui évolue aléatoirement entre les états s_1, s_2, s_3 et s_4 selon la matrice de probabilité de transition suivante :

$$P = \begin{pmatrix} 0.3 & 0.4 & 0.2 & 0.1 \\ 0.2 & 0.3 & 0.5 & 0.0 \\ 0.1 & 0.0 & 0.8 & 0.1 \\ 0.4 & 0.0 & 0.0 & 0.6 \end{pmatrix}$$

où le terme P_{ij} situé en ligne i et colonne j représente la probabilité de passer dans l'état s_j lorsqu'on est dans l'état s_i.
À chaque pas de temps, le décideur peut soit quitter le système et recevoir une récompense $R = 5$ unités, soit rester dans le système et recevoir une récompense $r(s_i)$ unités si le système est dans l'état s_i. Si le décideur choisit de rester dans le système, son état au pas de décision suivant sera déterminé par la matrice P. Dans la suite de l'exercice, on supposera que $r(s_i) = i$, $i = 1, \ldots, 4$ et que le taux d'actualisation des gains futurs est 0.5 et on va s'intéresser la recherche d'une politique optimale à l'horizon infini.

Q 1.1 Formuler ce problème comme un processus décisionnel Markovien. On précisera les états, les transitions et les récompenses.

Q 1.2 Soient v_i la valeur en l'état s_i, $i = 1, \ldots, 4$ d'une politique optimale pour le problème ci-dessus. Ecrire les équations de Bellman en fonction des variables $v_i, i = \ldots, 4$.

Q 1.3 Formuler un programme linéaire permettant de déterminer une politique optimale pour le problème.

Q 1.4 Executer les 3 premières itérations de l'algorithme d'itération de la valeur. Sachant que cet algorithme, effectué à une précision de 1/100, se termine par les itérations suivantes, en déduire la politique optimale :

itér.	s_1	s_2	s_3	s_4
0	0.00	0.00	0.00	0.00
.	...	...	...	...
6	5.00	5.00	5.97	7.13
7	5.00	5.00	5.99	7.14
8	5.00	5.00	6.00	7.14

Q 1.5 En partant de la politique qui consiste à quitter le système quelque soit l'état observé, appliquer l'algorithme d'itération de la politique pour retrouver le résultat de la question précédente.

Q 1.6 On s'intéresse maintenant à la politique π consistant à rester dans le système tant que l'on n'a pas rencontré l'état s_4 puis de sortir. Après avoir montré que cette politique n'est pas optimale, proposer une méthode permettant de déterminer la valeur minimale de R pour laquelle π est optimale.

Exercice 2 (3 points – Décision avec un critère RDU)

Q 2.1 Soit $\mathcal{X} = \mathbb{R}$ un ensemble de conséquences. La relation de préférences $\succsim_{\mathcal{X}}$ du décideur sur $\mathcal{X}$ coïncide avec la relation $\geq$. Soient deux loteries $P = \langle(x_1, p_1); \dots ; (x_n, p_n)\rangle$ et $Q = \langle(y_1, p_1); \dots ; (y_n, p_n)\rangle$ où les x_i et les y_i sont triés par ordre croissant de preférence sur l'espace des conséquences. Autrement dit, $x_1 < x_2 < \dots < x_n$ et $y_1 < y_2 < \dots < y_n$. Supposons qu'il existe $i_0 \in \{1, \dots , n\}$ tel que $x_i = y_i$ pour tout $i \neq i_0$, et tel que $y_{i_0} \in]x_{i_0}, x_{i_0+1}[$ si $i_0 < n$ et $y_{i_0} \in]x_{i_0}, +\infty[$ sinon. Montrez que, selon le critère RDU, la loterie Q est préférée à P.

Q 2.2 Soient deux loteries $P = \langle(x_1, p_1); \dots ; (x_n, p_n)\rangle$ et $Q = \langle(x_1, q_1); \dots ; (x_n, q_n)\rangle$ où les x_i sont triés par ordre croissant de preférence. Supposons qu'il existe $i_0 \in \{1, \dots , n-1\}$ tel que $p_i = q_i$ pour tout $i \notin \{i_0, i_0 + 1\}$, et tel que $q_{i_0} = p_{i_0} - \alpha$ et $q_{i_0+1} = p_{i_0+1} + \alpha$, pour un α positif. Montrez que, selon le critère RDU, la loterie Q est préférée à P.

Q 2.3 Soient deux loteries $P = \langle(x_1, p_1); \dots ; (x_n, p_n)\rangle$ et $Q = \langle(y_1, q_1); \dots ; (y_n, q_n)\rangle$ où les x_i et les y_i sont triés par ordre croissant de preférence. Supposons que, pour tout i, $y_i \geq x_i$ et que $\sum_{k=i}^{n} q_k \geq \sum_{k=i}^{n} p_k$. Montrez que, selon le critère RDU, la loterie Q est préférée à P.

Exercice 3 (4 points – Fonctions de croyance)

Un bouquiniste vend des romans (R) et des bandes dessinées (B) sur les quais de Seine. Ces livres peuvent être soit neufs (N), soit d'occasion (O). Dans son inventaire, il sait avec certitude qu'il possède 10% de romans neufs, 20% de romans d'occasion, 30% de bandes dessinées neuves et 20% de bandes dessinées d'occasion. Le bouquiniste sait qu'il a en plus 10% de romans dont il ne sait s'ils sont neufs ou d'occasion, et 10% de livres neufs dont il ne sait si ce sont des romans ou des bandes dessinées.

Q 3.1 Donnez une expression de l'ensemble des lois de probabilités $\mathcal{P}$ compatibles avec les informations ci-dessus.

Ensemble des événements élémentaires : $\mathcal{X} = \{RO, RN, BO, BC\}$.
$\mathcal{P} = \{P(RO) = \alpha, P(RN) = \beta, P(BO) = \gamma, P(BN) = \delta :$
$\alpha = 0,2 + t,$
$\beta = 0,1 + y + z,$
$\gamma = 0,2,$
$\delta = 0,3 + x,$
$x + y = 0,1$
$z + t = 0,1\}.$

Q 3.2 Calculez l'enveloppe inférieure f de $\mathcal{P}$ et montrez que $\mathcal{P} = \{$ lois de probas $P : P \geq f\}$.

	$\emptyset$	RO	RN	BO	BN	R	O	$RO \cup BN$	$RN \cup BO$	N	B	$R \cup BO$	$R \cup BN$	$O \cup BN$	$N \cup BO$	$\mathcal{X}$
ϕ	0	0.2	0.1	0.2	0.3	0.1	0	0	0	0.1	0	0	0	0	0	0
f	0	0.2	0.1	0.2	0.3	0.4	0.4	0.5	0.3	0.5	0.5	0.6	0.8	0.7	0.7	1

Q 3.3 Calculez l'inverse de Möbius ϕ de f. Montrez que f est une fonction de croyance.

$$\left.\begin{array}{r} \phi(\emptyset) = 0 \\ \displaystyle\sum_{B \subseteq \mathcal{X}} \phi(B) = 1 \\ \forall B \subseteq \mathcal{X},\ \phi(B) \geq 0 \end{array}\right\} \Longrightarrow f \text{ est une fonction de croyance.}$$

Exercice 4 (3 points – Utilités qualitatives)

Q 4.1 Soit π la distribution de possibilité caractérisée sur les états de la nature par $\pi(s_1) = 1$ et $\pi(s_2) = 0.4$. Soit X un ensemble de conséquences contenant au moins trois éléments x_1, x_2, x_3 distincts. On considère alors les actes suivants :

	s_1	s_2
f	x_2	x_1
g	x_2	x_3

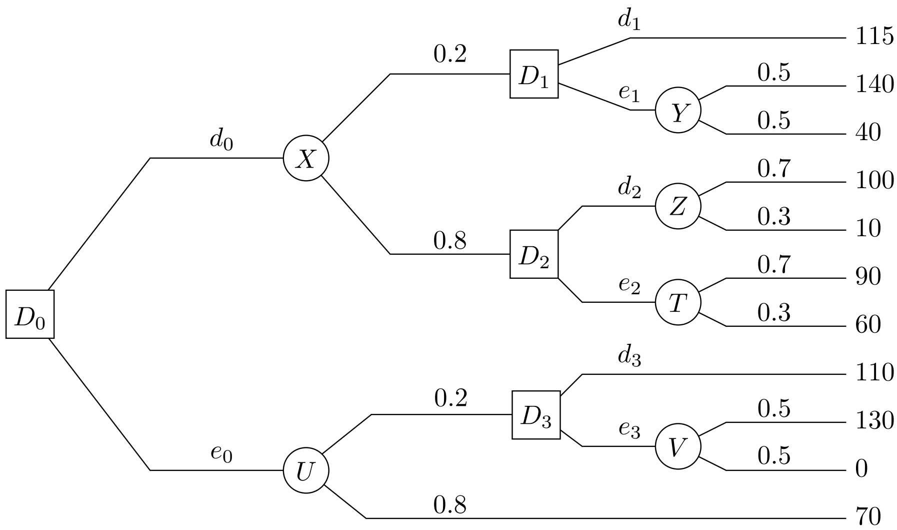
avec $u(x_1) = 0.8, u(x_2) = 0.6, u(x_3) = 0.2$. Sur la base de ces données, expliquez pourquoi on devrait préférer f à g. Montrez que cette préférence n'est pas compatible avec le critère de l'utilité optimiste ni avec celui de l'utilité pessimiste.

Q 4.2 À partir de u on construit l'utilité binaire u_2 définie pour toute conséquence x par :
$u_2(x) = \langle 1,\ 2 - 2u(x)\rangle$ si $u(x) \geq 1/2$ et $u_2(x) = \langle 2u(x),\ 1\rangle$ sinon.
Après avoir montré que u et u_2 induisent le même ordre sur les conséquences de X, comparer les actes f et g avec le critère de l'utilité binaire possibiliste de Giang et Shenoy. Que peut-on observer ?

Q 4.3 Le critère de l'utilité binaire possibiliste vérifie t-il le principe de la chose sûre de Savage (sure thing principle) ? Dans l'affirmative on fera la démonstration, dans la négative on donnera un contre-exemple.

Exercice 5 (4 points – Décision avec un critère RDU)

On considère l'arbre de décision ci-dessous. Sur les feuilles sont indiquées les utilités du décideur.



Q 5.1 Supposons que le décideur utilise un critère RDU avec pour fonction de déformation des probabilités φ la fonction affine par morceaux dont les segments sont définis par :

p	$\varphi(p)$	p	$\varphi(p)$	p	$\varphi(p)$	p	$\varphi(p)$	p	$\varphi(p)$
0	0	0,1	0,3	0,2	0,4	0,3	0,55	0,4	0,65
0,46	0,68	0,5	0,7	0,56	0,75	0,6	0,77	0,66	0,8
0,7	0,84	0,76	0,87	0,8	0,9	0,9	0,95	1	1

Q 5.1.1 Déterminer la valeur de RDU aux nœuds D_1, D_2 et D_3.

$\text{RDU}(d_1) = 115$
$\text{RDU}(e_1) = 40 + \varphi(0.5)[140 - 40] = 110$
$\text{RDU}(d_2) = 10 + \varphi(0.7)[100 - 10] = 85.6$
$\text{RDU}(e_2) = 60 + \varphi(0.7)[90 - 60] = 85.2$
$\text{RDU}(d_3) = 110$
$\text{RDU}(e_3) = 0 + \varphi(0.5)[130 - 0] = 91$

Q 5.1.2 Déterminer la valeur de RDU pour l'ensemble des stratégies à la racine D_0.

$\text{RDU}(d_0, d_1, d_2) = 10 + \varphi(0.76)[100 - 10] + \varphi(0.2)[115 - 100] = 94.3$
$\text{RDU}(d_0, d_1, e_2) = 60 + \varphi(0.76)[90 - 60] + \varphi(0.2)[115 - 90] = 96.1$
$\text{RDU}(d_0, e_1, d_2) = 10 + \varphi(0.76)[40 - 10] + \varphi(0.66)[100 - 40] + \varphi(0.1)[140 - 100] = 96.1$
$\text{RDU}(d_0, e_1, e_2) = 40 + \varphi(0.9)[60 - 40] + \varphi(0.66)[90 - 60] + \varphi(0.1)[140 - 90] = 98$
$\text{RDU}(e_0, d_3) = 70 + \varphi(0.2)[110 - 70] = 86$
$\text{RDU}(e_0, e_3) = 0 + \varphi(0.9)[70 - 0] + \varphi(0.1)[130 - 70] = 84.5$

Q 5.1.3 Déterminer la stratégie optimale selon RDU.

$\Longrightarrow$ stratégie gagnante : (d_0, e_1, e_2)